Introducción a la ética

Colección Teorema

Bernard Williams

Introducción a la ética

CUARTA EDICIÓN

CÁTEDRA
TEOREMA

Título original de la obra: *Morality: An Introduction to Ethics*
Traducción de Manuel Jiménez Redondo

Reservados todos los derechos. El contenido de esta obra está protegido por la Ley, que establece penas de prisión y/o multas, además de las correspondientes indemnizaciones por daños y perjuicios para quienes reprodujeren, plagiaren, distribuyeren o comunicaren públicamente, en todo o en parte, una obra literaria, artística o científica, o su transformación, interpretación o ejecución artística fijada en cualquier tipo de soporte o comunicada a través de cualquier medio, sin la preceptiva autorización

© 1972, by Bernard Williams
Ediciones Cátedra, S. A., 1998
Juan Ignacio Luca de Tena, 15. 28027 Madrid
Depósito legal: M. 42.669/1998
ISBN: 84-376-0337-4
Printed in Spain
Impreso en Fernández Ciudad, S. L. Catalina Suárez, 19
28007 Madrid

Índice

Prefacio 11
El hombre amoral 17
Subjetivismo: consideraciones iniciales 27
Interludio: relativismo 33
Subjetivismo: otras consideraciones 39
«Bueno» 51
Bondad y roles 61
Estándares morales y la nota distintiva del hombre . . 69
Dios, moralidad y prudencia 77
¿Sobre qué versa la moralidad? 87
Utilitarismo 95

A mi madre y mi padre

Prefacio

Diríase que escribir sobre filosofía moral es un asunto arriesgado, y no precisamente por las dificultades inherentes a escribir sobre cualquier materia complicada, o sobre cualquier materia en general, sino por dos razones especiales. La primera es que es probable que uno ponga de manifiesto las limitaciones y la inadecuación de la visión que uno tiene del asunto más directamente, que, por lo menos, en otras partes de la filosofía. La segunda es que uno podría correr el riesgo, si se le toma en serio, de extraviar a la gente en asuntos que son de importancia. Mientras que son pocos los tratadistas de esta materia que hayan sido capaces de evitar el primer riesgo, son muchos los que han evitado el segundo o bien haciendo imposible que se les tome en serio, o bien rehuyendo escribir sobre algo de importancia, o bien por ambos medios.

Esta triste verdad se aduce con frecuencia como una acusación especial contra la filosofía moral contemporánea de estilo «analítico» o «lingüístico»: que es una filosofía vacía y aburrida. En un cierto sentido, como acusación especial, la objeción no es equitativa: casi toda la filosofía moral de casi todas las épocas ha sido vacía y aburrida, y el número de grandes obras en la materia (en comparación con el número de libros que de una u otra forma versan sobre la moralidad) puede literalmente contarse con los dedos de una mano. La vaciedad de las obras del pasado, sin embargo, ha sido con frecuencia la vaciedad del moralizar convencional, el tratamiento

banal de los asuntos morales. La filosofía moral contemporánea ha encontrado una forma original de ser aburrida, que no es otra que el no hablar de asuntos morales en absoluto. O mejor, no se trata tanto de que se haya desarrollado un estilo de filosofía moral que corta por completo toda conexión con asuntos morales —esto, si fuera posible, tendría el interés de ser cosa notable—, sino de que el deseo de reducir al mínimo el compromiso moral manifiesto y de usar argumentos morales tan sólo a título de irrefragablemente ilustrativos deja la impresión de que todos los asuntos morales importantes han quedado fuera, en alguna parte, y que se ha hecho uso de gran cautela y poca imaginación al dejar aparecer aspectos menores de algunos de ellos.

Existen muchas razones que explican esta situación. Una razón central es que los puntos de vista contemporáneos sobre la moralidad no dicen nada claro sobre cuáles son las cualidades mentales o de carácter más adecuadas para el pensamiento moral constructivo (por supuesto que en algunas explicaciones de la moralidad ni siquiera queda claro que pueda existir una cosa tal como pensamiento moral constructivo); de ahí que lo dejen a uno inquieto sin saber si es probable que esas cualidades sean cualidades de filósofos, cuando la filosofía es en buena parte una actividad profesional y académica que exige principalmente, aunque no exclusivamente, habilidades discursivas y analíticas. Si existiera una actividad consistente en deducir conclusiones morales sustantivas de premisas *a priori*, podría razonablemente esperarse que filósofos bien entrenados fueran especialmente aptos para ella; pero no existe, y el hecho de que si existiera, los filósofos profesionales tendrían una buena razón para suponérseles informados sobre moralidad, es una de las buenas razones para pensar que no existe tal actividad.

Ciertamente que el problema no es, como alguno pretende, que si el filósofo no se distancia de forma patente e incluso si no es metodológico, entonces tiene que estar *predicando;* posiblemente no pueda ser esta la única alternativa. Se trata más bien de un problema de estilo,

en el sentido más profundo de «estilo», en el que descubrir el estilo correcto es descubrir qué es lo que uno está realmente tratando de hacer. ¿Cómo combinar la argumentación (que probablemente sea, después de todo, el único derecho con que cuenta el filósofo sobre la atención de la gente) con vuelos más largos o detalles más concretos, de forma que el tejido de pensamiento moral resulte lo más interesante posible? ¿Puede la realidad de las situaciones morales complejas ser representada por otros medios que por los de la literatura imaginativa? Y si no, ¿pueden otros acercamientos más esquemáticos representar suficientemente a la realidad? ¿Cuánto de lo que genuinamente preocupa a la gente es susceptible de teoría general?

Si conociera las respuestas a estas preguntas no tendría que hacerlas ahora. Este ensayo sigue un curso tortuoso, y aunque he tratado de señalizar los principales recodos, tal vez sea interesante esbozar un plan de antemano. Empiezo con una figura, que con frecuencia ha interesado y por supuesto también ha preocupado a los moralistas porque representa un desafío para la moralidad y también una exigencia de que se la justifique: el hombre amoral, al que se supone inmune a las consideraciones morales. Algunas de las cuestiones más interesantes sobre él, que yo apenas he tocado, radican no tanto en lo que se le pudiera decir a él como en lo que se puede decir sobre él: qué clase de consistencia puede ofrecer en ese respecto el hombre amoral. De él pasamos al que no rechaza la moralidad, pero que adopta sobre su naturaleza un punto de vista especial y casi tan inquietante como el anterior: subjetivistas de varias clases, y una clase desvergonzadamente crasa (pero común) de relativismo. En ese punto examino con detalle un proyecto muy acariciado por buena parte de la filosofía moral moderna, al que he llamado subjetivismo desactivado.

De ahí paso a algunas consideraciones sobre la bondad; y, tratando de arrojar alguna claridad sobre las formas en que un hombre puede ser bueno *en* ciertas cosas, y más aún sobre las cosas como las que puede ser bueno

un hombre, trato de deslindar algunas cuestiones puramente lógicas de lo que me parecen asuntos más sustantivos sobre lo que los hombres son y sobre la conexión de todo eso con la bondad. Dos cuestiones en particular sobresalen entre las muchas que se amontonan en torno a este área: las relaciones de la realización intelectual con los estándares de moralidad, y la cuestión de si la existencia de Dios supondría alguna diferencia en la situación de la moralidad. Esto suscita algunas cuestiones generales importantes sobre motivos morales y no morales. Estas conducen a su vez a algunos asuntos relativos al núcleo o sustancia de la moralidad, y a la cuestión de si, en última instancia, de lo único de lo que en la moralidad se trata es del bienestar o (de forma más restringida, quizá) de la felicidad del hombre. Finalmente se aborda la forma menos complicada de hacer apuntar la moralidad a la felicidad, la del utilitarismo, pero sólo con la extensión precisa para sugerir qué clase o sistema tan especial y peculiar es el utilitarismo cuando se entiende adecuadamente; y para indicar la dirección en la que hay que buscar sus peculiaridades. Desarrollarlas es tarea que ha de quedar para otra ocasión.

Una de las razones por las que este libro no puede ser considerado como un libro de texto, ni tan siquiera un libro de texto introductorio, o como un bosquejo de filosofía moral, es que no ofrece una teoría sistemática. No me importa confesarlo así, ya que me parece que esta materia, a pesar de que se presta mucho menos a ello, ha sido objeto de más sistematizaciones supergenerales y supersimplificadas que, virtualmente, cualquier otra parte de la filosofía. Con esto no quiero decir que hubiera que acercarse a la filosofía moral sin preconcepciones (lo que sería imposible) o incluso sin preconcepciones teóricas (lo que muy bien pudiera resultar estático y estéril). Se trata simplemente de que, para empezar, de lo que uno debiera ocuparse es de los fenómenos morales, tal como se le ofrecen en su propia experiencia e imaginación, y en un nivel más teórico, de las exigencias de las *otras* partes de la filosofía, en particular de la filosofía de la mente. No hay razón para

suponer que la filosofía moral, o algo, a veces más amplio y a veces más restringido, llamado «teoría del valor», tuviera que arrojar en absoluto una teoría autosuficiente.

Otra razón por la que éste no es un libro de texto es que deja fuera amplios tramos de la materia. Esto es por lo menos completamente obvio. Pero para que lo que digo quede en una mejor perspectiva no está de más mencionar un par de temas que en un planteamiento más amplio de la filosofía moral tendrían que ocupar a mi entender un puesto central: qué son pensamiento práctico y actuar por una razón; qué es consistencia en la acción y consistencia en el pensamiento moral y, en relación con ello, por qué el conflicto moral es un hecho básico de la moralidad; por qué la noción de *regla* es importante para alguna de las partes de la moralidad, no para todas (el presente libro nada tiene que decir sobre su importancia); lo inconstante y problemática que es la distinción entre «moral» y «no moral», sobre todo en su uso más importante: cuando se trata de distinguir entre distintas clases de excelencia humana.

Que este ensayo dejara fuera muchas cosas de importancia era inevitable; que siguiera un camino tortuoso, no lo era. A lo que no sabría ciertamente responder es a la pregunta de si era inevitable que no lograra dar con una solución al problema de cómo escribir sobre filosofía moral.

El hombre amoral

«¿Por qué habría yo de hacer algo?» Dos de las muchas formas en que se puede abordar esta cuestión son éstas: como una expresión de desesperación, cuando significa algo así como «dame una razón para hacer algo; nada tiene sentido»; y como una expresión de tono más desafiante, contra la moralidad, cuando significa algo así como «¿por qué hay algo que yo *tuviera que*, *debiera* hacer?».

Aun cuando a la cuestión, en su primer sentido, podríamos parafrasearla como «dame una razón para...», no está nada claro que de hecho pudiéramos dar al hombre que hace esa pregunta una razón, no está nada claro que empezando de tan abajo pudiéramos inyectarle, por medio de *argumentos*, interés por algo. Podríamos, por supuesto, «darle una razón» en el sentido de encontrarle algo por lo que está dispuesto a interesarse, pero esto no es inducirle a interesarse por vía de razonamiento, y es muy dudoso que pueda existir tal cosa. Lo que necesita es ayuda, esperanza, no razonamientos. Por supuesto que si sigue viviendo, estará haciendo *alguna cosa* y no otra y que, por tanto, en algún sentido absolutamente mínimo tiene alguna clase de razón, alguna preferencia mínima por esas cosas y no por otras. Pero subrayar esto difícilmente nos llevará a ninguna parte, pues hace esas cosas mecánicamente tal vez, para seguir tirando, sin que por ello signifiquen nada para él. A su vez, si considera su estado como una razón para el suicidio, entonces eso será tomar una decisión efectiva; como for-

ma de escapar a la toma de toda decisión, el suicidio, como decisión que es, llega inevitablemente demasiado tarde (como subraya Camus en *Le Mythe de Sisyphe)*. Pero no supondría ninguna victoria para nosotros ni para él si resultara que, después de todo, había justamente una decisión que estaba dispuesto a reconocer, precisamente ésa.

No veo por qué tendría que considerarse como una derrota de la razón o de la racionalidad la impotencia frente al estado de este hombre; su estado representa más bien una derrota para la humanidad. Pero en relación con el hombre que hace la pregunta en el segundo sentido, muchos moralistas han considerado que representa un desafío efectivo para el razonamiento moral. Pues este hombre reconoce, después de todo, algunas razones para hacer cosas; es además, como la mayoría de nosotros, alguien de nuestro tiempo. Si la moralidad puede levantar el vuelo racionalmente, tendrá que ser capaz de hacerlo con un argumento contra él; pues, aun cuando en su forma pura —en la que podemos llamarle el *hombre amoral*—, tal vez no pueda ser persuadido de hecho, parece que sí sería un respiro para la moralidad el que existieran razones capaces de persuadir a ese hombre si fuera racional.

Podemos empezar preguntando qué clase de motivaciones tiene. Ese hombre es indiferente a consideraciones morales, pero existen cosas por las que se preocupa, y tiene algunas preferencias y fines reales. Puede que, presumiblemente, éstos sean el placer o el poder; o puede que se trate de algo menos común, como la pasión por coleccionar cosas. Ahora bien, estos fines, por sí mismos, no excluyen algún reconocimiento de la moralidad. ¿Qué es lo que tenemos que dejar fuera para representárnoslo como un hombre que efectivamente se niega a prestar ese reconocimiento? Presumiblemente cosas tales como preocuparse por los intereses de otras personas, tener alguna propensión a decir la verdad o a guardar las promesas incluso cuando no le viene bien, estar dispuesto a rechazar formas de acción en razón de que no son equitativas, o son deshonestas o son egoístas. Estos

son algunos de los aspectos sustantivos de la moralidad. Tendríamos quizás también que dejar fuera un aspecto más formal de la moralidad, a saber, cualquier disposición por su parte a recapacitar y dar en la idea de que si está bien para él actuar de esa forma, no puede reprocharse a los otros el que actúen de forma similar contra él. Pues si está dispuesto a adoptar esta actitud, nosotros, por nuestra parte, seríamos capaces de dar un paso adelante diciendo que no se trataba de un hombre sin moralidad, sino de un hombre con una moralidad peculiar.

Sin embargo, aquí necesitamos hacer una distinción. En un cierto sentido es posible que un hombre piense «me parece muy bien» que cada uno se comporte de forma autointeresada, sin que por ello haya penetrado en absoluto en un terreno de pensamiento que quepa calificar de distintivamente moral: este es el caso cuando «me parece muy bien» viene más o menos a significar «no me voy a poner a moralizar sobre eso». Estará en algún terreno moral si «me parece muy bien» significa algo así como «permitido», pues esto llevaría a implicaciones tales como «la gente no debe interferirse cuando los demás van en pos de sus propios intereses», y esto no es una idea que pueda corresponderle como al hombre amoral que es. Igualmente, si protesta (como sin duda hará) de que otros le traten como él los trata, esto será perfectamente coherente mientras su protesta consista en cosas tales como manifestaciones de disgusto o en contraatacar a su vez. Lo que no puede hacer de forma coherente es *ofenderse* por ello o desaprobarlo, pues éstas son actitudes que caen dentro del sistema moral. Puede que resulte difícil averiguar si en realidad ha hecho entrega o no de este rehén a la argumentación moral, pues sin duda habrá descubierto que las expresiones insinceras de desaprobación y lesión moral sirven para disuadir de acciones hostiles a algunos de los miembros más aprensivos de su entorno.

Esto ilustra, como lo hacen muchas de sus actividades, el hecho obvio de que este hombre es un parásito del sistema moral, y que ni él ni sus satisfacciones podrían

existir en la forma en que lo hacen si los otros no operaran de forma diferente. Pues, en general, no puede haber sociedad sin algunas reglas morales, y él necesita de la sociedad; también obtiene muy particulares ventajas de instituciones morales tales como la promesa y de las disposiciones morales de la gente que le rodea. No puede negar, como cuestión de hecho, su postura parasitaria; aunque sí que se mostrará muy resistente a las sugerencias de su relevancia. Pues si tratamos de decirle: «¿Qué sería de ti si todos se comportaran como tú?», responderá: «Bien, si lo hicieran, nada bueno, supongo; aunque de hecho tal vez podría obtener más ventajas en el caos resultante que algunos de ellos. Pero el hecho es que la mayoría de ellos no lo va a hacer; y si alguna vez por fin se decidieran, para entonces yo ya estaré muerto.» La apelación a las consecuencias de una universalización hipotética es un argumento esencialmente moral, y en consecuencia, y coherentemente, nuestro hombre amoral no se deja impresionar por esa apelación.

Para mantener esta actitud, existen bastantes cosas que tiene que evitar si es que quiere ser consistente. Una —como ya hemos indicado antes— es cualquier tendencia a decir que la mayoría más o menos moral *no tiene derecho* a marginarlo, rechazarlo o tratarlo como enemigo, en el caso de que esa mayoría se sienta inclinada a hacerlo así (su poder, o encanto personal, o capacidad de disimulo pueden ser tales que no sea ese el caso), pues, como el amoral que es, no puede pensar en justificaciones, por lo menos de esa clase. Además, si quiere ser coherente, tiene que resistir a la tendencia más insidiosa de atribuirse a sí mismo una forma de ser realmente espléndida y en particular el considerarse notablemente valiente en comparación con la cobarde multitud. Pues acariciando tales pensamientos correrá un constante peligro de salirse fuera del mundo de sus propios deseos y gustos, y de penetrar en una región en la que la posesión de ciertas disposiciones es considerada como una excelencia en los seres humanos o como un bien para la sociedad, o cosas de este tipo; y si bien es verdad que tales pensamientos no tienen por qué conducir directamente a considera-

ciones morales, sí que representan un acercamiento sustancial a ellas, pues inmediatamente llevan a preguntas sobre qué es lo bueno en esas disposiciones, y le resultará muy difícil ir muy lejos en esas cuestiones sin pensar en términos de las necesidades e intereses *generales* del prójimo, lo que le retrotraería al mundo del pensamiento moral del que se está excluyendo a sí mismo.

La tentación de considerarse a sí mismo valiente es particularmente peligrosa, pues se trata casi de una noción moral y comporta toda una serie de reflexiones distintivamente morales. La aplicación que este hombre haga de la noción tendrá también una presuposición que es falsa, a saber: que los ciudadanos más morales serían amorales si pudieran hacerlo impunemente, si no tuvieran tanto miedo, o si no estuvieran pasivamente condicionados por la sociedad —si, en general, no padecieran de inhibiciones. Es la idea de que tienen miedo lo que da alas a la estimación de su propio coraje. Pero estas presuposiciones son absurdas. Si lo que quiere decir es que, si como individuo uno estuviera seguro de poderlo hacer impunemente, uno rompería con toda regla moral (la idea que subyace al modelo de la invisibilidad del anillo de Giges en *La República*, de Platón), esto es falso de muchos agentes, y la razón es la siguiente: las reglas y conceptos morales más básicos quedan fuertemente internalizados en el proceso de socialización, y a partir de un determinado nivel de ese proceso no se evaporan sin más porque desaparezca la policía o la censura del prójimo. Y esto es ingrediente constitutivo de su carácter de reglas morales, por oposición a los requerimientos *meramente* legales o a los asuntos de convención social. Los efectos de la educación moral pueden hacer que la gente *quiera* actuar, con mucha frecuencia, de forma no auto-interesada, y, por lo menos, muchas veces logran hacer muy difícil, por razones internas, el comportamiento detestable.

Pero esto, dirá nuestro hombre, no es más que condicionamiento social; quitad eso, y no hallaréis motivaciones morales. Podemos rechazar la retórica del término «condicionamiento»; incluso si hubiera una teoría

verdadera, que no la hay, que pudiera explicar toda la educación moral y similar en términos de la teoría behaviourista del comportamiento, tendría que explicar todavía las muy evidentes diferencias que existen entre una socialización inteligente y lograda, con el resultado de la capacidad de habérselas consigo mismo y con las cosas, y la producción de reflejos condicionados. Pero digamos en lugar de eso que toda motivación moral es producto de influencias sociales, de la enseñanza, de la cultura, etc. Sin duda que es verdad. Pero virtualmente cualquier otra cosa que se refiera al hombre es producto de lo mismo, incluyendo su lenguaje, sus métodos de pensamiento, sus gustos e incluso sus emociones, incluyendo la mayor parte de las disposiciones de las que hace provisión el hombre amoral. Pero, puede replicar, supongamos que admitimos que cualquier complejo, incluso mis deseos, están influidos por la cultura y el entorno, y en muchos casos producidos por ellos; sin embargo, existen impulsos *elementales*, de una clase autointeresada, que están en la base de todo ello: son esos impulsos los que constituyen lo que los hombres *realmente* son.

Si «básico» significa «genéticamente primitivo», es posible que tenga razón: es un asunto de teoría psicológica. Pero, aunque verdadera en este sentido, su afirmación resulta irrelevante una vez más (para su argumentación, no para cuestiones relativas a cómo educar a los niños); si existe algo así como lo que los hombres *realmente* son, eso no será idéntico a lo que son los niños pequeños, porque los niños pequeños no tienen lenguaje ni tampoco otras muchas cosas que los hombres realmente tienen. Y si, a su vez, la prueba de lo que los hombres *realmente* son se hace viendo cómo los hombres se comportan en condiciones de una gran tensión, privación o carencia (el *test* que impuso Hobbes en su descripción del estado de naturaleza), lo único que uno puede hacer es volver a preguntar por qué tiene que ser ése el *test*. Aparte de lo incierto de sus resultados, ¿por qué habría de ser ése el test idóneo? Condiciones de gran tensión y privación no son las condiciones para

observar la conducta típica de ningún animal ni para observar otras características de los seres humanos. Si alguien dice que si deseamos saber lo que los hombres *realmente* son, los observemos después de que hayan pasado tres semanas en un bote salvavidas, no está claro por qué ha de ser esto una mejor máxima en lo concerniente a sus motivaciones que en lo concerniente a su condición física.

Si existe algo así como lo que los hombres *realmente* son, bien puede ser que (por lo menos en esta clase de aspectos) ello no sea tan diferente de lo que los hombres *de hecho* son; esto es, seres en cuyas vidas las consideraciones morales juegan un papel importante, formativo, pero a veces inseguro.

El hombre amoral, pues, haría bien en evitar la mayoría de las formas de comparación autocongratulatoria de sí mismo con el resto de la sociedad. El resto puede, por supuesto, tener una cierta tendencia a admirarlo, o si no el resto, sí aquellos que se encuentran a una distancia suficiente como para que nuestro hombre no interfiera de forma directa en sus intereses y afecciones. Empero, no por ello debería sentirse demasiado alentado, pues probablemente se trata de la satisfacción sustitutiva de una apetencia (lo que no significa que serían como él si pudieran, pues una veleidad es cosa distinta de un deseo frustrado). Y no lo admirarán ni les resultará grato si, por lo menos, no es reconociblemente humano. Y esto suscita la cuestión de si le hemos dejado lo suficiente para ser eso.

¿Se preocupa por alguien? ¿Existe alguien cuyos sufrimientos y afecciones le afecten? Si a esto respondemos con una negativa, se diría que lo que hemos descrito es un psicópata. Y si se trata de un psicópata, la idea de argumentar con él en favor de la moralidad puede ser ciertamente vana, pero el hecho de que sea vana no tiene por qué socavar la base de la moralidad o de la racionalidad. Pues la actividad de justificar la moralidad ha de extraer el interés que tiene de la existencia de una alternativa —de que haya algo *contra* lo que justificarla. El hombre amoral parecía importante porque parecía re-

presentar esta alternativa; su vida, después de todo, parecía no carecer de atractivos. El psicópata es, en cierto sentido, importante para el pensamiento moral; pero su importancia radica en el hecho de que nos horroriza, y nos vemos en la necesidad de encontrar una explicación más profunda de cómo y por qué nos horroriza. Su importancia no radica en que tenga atractivo como forma alternativa de vida.

El hombre amoral, cuya figura hemos bosquejado antes a grandes trazos, parecía que era posible que tuviese algún atractivo más que ése; uno podría imaginárselo con afectos y preocupado a veces por lo que ocurre a otros. A uno le viene a la mente algún estereotipo de una película de gangsters, de esa figura despiadada y brillante que se preocupa por su madre, por su hijo, o incluso por su querida. Es todavía reconociblemente amoral en el sentido de que para él no pesan las consideraciones generales, y es extremadamente corto en consideraciones de imparcialidad o de tipo similar. Aunque de cuando en cuando se preocupa por otras personas, todo ello depende del humor en que se encuentre. Con este hombre, por supuesto, de muy poco van a valer los argumentos de filosofía moral, aunque sólo sea por el hecho de que, a su juicio, siempre va a tener algo mejor que hacer que dedicarse a escucharlos. Esta no es la cuestión (aunque sí más cuestión que lo que algunas discusiones sobre la argumentación moral nos hacen suponer). La cuestión es más bien que nuestro hombre nos proporciona un modelo, en términos del cual nos podemos hacer una idea de lo que las consideraciones morales necesitan para remontar el vuelo, aunque no sea probable que en la práctica puedan remontarlo en una conversación con él.

Nos proporciona, creo yo, casi lo suficiente. Nuestro hombre tiene la idea de hacer algo *por* alguien, porque esa persona necesita algo. En realidad, sólo opera con esta noción cuando se siente inclinado a ello; pero la idea misma no es la de sentirse inclinado a ello. Incluso si ayuda a esas personas porque le viene en gana o porque le resultan gratas, y no por otra razón (y no es

que, por lo que respecta a esas acciones particulares, necesite perfeccionar esas excelentes razones), lo que desea hacer es *ayudarles en su necesidad*, y la idea que tiene cuando le gusta alguien y actúa de esa forma es «necesita ayuda», y no «me gusta y necesita ayuda». Esto es una cuestión vital: este hombre es capaz de pensar en términos de los intereses de los otros, y su incapacidad para ser un agente moral radica (en parte) en que sólo está dispuesto a actuar así de forma intermitente y caprichosa. Pero no existe ninguna cima insalvable entre su estado y las disposiciones básicas de la moralidad. Hay personas que necesitan ayuda, que por el momento no son personas a las que él tenga ganas de ayudar o le resulte grato ayudar; y hay otras personas que quieren y desean socorrer a otras personas particulares en necesidad. Conseguir que considere la situación de esas personas parece que ha de consistir más bien en una extensión de su imaginación y su comprensión que en salto a otra cosa muy distinta, el «plano moral». Y si consiguiéramos que considerara la situación de esas personas, en el sentido de pensar en ella e imaginársela, es concebible que empezara a mostrar alguna consideración por ella: estamos extendiendo sus simpatías. Y si podemos conseguir extender sus simpatías a personas inmediatas que necesitan su ayuda, tal vez fuéramos capaces de hacerlo también para personas menos inmediatas cuyos intereses han sido violados, logrando así que llegara a una comprensión, aunque sólo fuera primitiva, de las nociones de justicia. Si lográramos que recorriera todo este camino, si bien es verdad que su capacidad para ulteriores consideraciones morales sigue siendo extremadamente exigua, también lo es que alguna sí que tiene; en cualquiera de los casos, no es el hombre amoral del que habíamos partido.

La finalidad de este modelo no es la de bosquejar las líneas generales de una construcción de toda la dimensión moral a partir de la posibilidad de la simpatía y de las extensiones de la simpatía: eso sería imposible. (Ni Hume, que fue quien más se acercó a esa idea, lo intentó siquiera. Su sistema, entre las muchas cosas valiosas e interesantes

que contiene, tiene una distinción entre virtudes «naturales» y «artificiales» que es relevante para esta cuestión.) La finalidad del modelo es sugerir precisamente una cosa: que si partimos de un hombre que se preocupe, aunque sólo sea mínimamente, por los otros, entonces no tenemos que adscribirle ninguna clase fundamentalmente nueva de pensamiento o experiencia para incluirlo en el mundo de la moralidad, sino solamente lo que, a todas luces, es una extensión de lo que ya él tiene. Ciertamente que no ha penetrado muy lejos todavía y el territorio es muy extenso; como hemos visto al pintar al hombre amoral, hay que recorrer un largo camino para escapar de él. Pero el hombre con la simpatía así ampliada, con capacidad para pensar en las necesidades de la gente más allá de sus compromisos inmediatos, está reconociblemente en él.

De esto no se sigue que la preocupación simpatética por los otros sea una condición necesaria para estar en el mundo de la moralidad, que el camino bosquejado sea el *único* que nos introduzca en la moralidad. Eso no se sigue de lo dicho hasta el momento; pero es verdad.

Más tarde volveremos sobre las consideraciones que aquí hemos hecho sobre las motivaciones morales y otras clases de motivaciones. Ahora vamos a pasar a otra figura que también ha sido considerada perturbadora para la moralidad si bien de forma diferente que la del hombre amoral. Este hombre está contento de tener una moralidad, pero subraya que otras personas tienen una moralidad diferente —e insiste en que no hay forma de decidir entre ellas. Se trata del subjetivista.

Subjetivismo: consideraciones iniciales

Consideremos tres enunciados, cada uno de los cuales expresa a su manera la idea de que las opiniones morales, o los juicios morales, o las concepciones morales son «meramente subjetivos».

a) «Los juicios morales de un hombre meramente consignan (o expresan) sus propias actitudes.»

b) «De los juicios morales no se puede probar, decidir o mostrar que sean verdaderos como puede hacerse de los enunciados científicos; son cuestión de opinión individual.»

c) «No hay hechos morales; sólo hay la clase de hechos que la ciencia o la observación común pueden descubrir, y los valores que los hombres asignan a estos hechos.»

Estos juicios se aproximan mucho entre sí y en las discusiones sobre el subjetivismo y el objetivismo uno encuentra versiones de los tres que son usadas como virtualmente intercambiables. Por supuesto que la correlación entre estos juicios es genuina. Pero son significativamente diferentes. El primero, *a*), expresa lo que en un sentido amplio podríamos llamar un punto de vista *lógico* o lingüístico: pretende decirnos algo sobre lo que las observaciones morales son o hacen. El segundo, *b*), introduce un conjunto de nociones que no está presente en el primero, nociones conectadas con el concepto de *conocimiento*, y puede considerarse que expresa un punto de vista *epistemológico* sobre los juicios morales. El tercer enunciado, *c*), es el más vago y el menos

tangible de los tres, y muestra en su superficie el riesgo de hundirse, parcial o totalmente, en uno u otro de los dos primeros: que es lo que muchos filósofos pretenderían que debía hacer. Sin embargo, en su forma inadecuada parece hacer un ademán apuntando hacia algo sumamente próximo a aquello a lo que se han referido muchos a quienes ha preocupado la cuestión de la objetividad moral: la idea de que no existe un orden moral «ahí fuera» —ahí fuera, en el mundo, sólo existen las clases de cosas y las clases de hechos de los que trata la ciencia y otras formas más cotidianas de pesquisa humana de las que la ciencia constituye un refinamiento. *Alles anderes ist Menschenwelt.* Del enunciado *c*) puede decirse —usando el término sin ambiciones— que expresa un punto de vista *metafísico.*

El punto de vista metafísico saca a la luz, de la forma más explícita, algo que está latente en los tres enunciados, una distinción entre hecho y valor. Una preocupación central de buena parte de la filosofía moderna ha sido la distinción entre hecho y valor. Una importante forma que ha adoptado esta preocupación ha sido la de subrayar la distinción rechazando al mismo tiempo las consecuencias inquietantes que se suponía que dicha distinción había de tener, y ello tratando de mostrar o bien que las consecuencias no eran tales consecuencias o bien que no eran inquietantes. Este proyecto de subjetivismo *desactivado* (como muy bien podríamos llamarlo) puede ser expresado en términos de nuestros tres enunciados subjetivistas aproximadamente como sigue: que, en la medida en que son defendibles, vienen a decir lo mismo; y que lo que dicen, primero, no es alarmante y, segundo, es esencial para la naturaleza de la moralidad (a veces se considera, y de forma bien extraña por cierto, que el que sea esencial para la naturaleza de la moralidad implica por sí mismo el que no sea alarmante).

En lo que sigue nos vamos a dedicar a examinar este proyecto, haciendo alguna interrupción. Empieza como sigue. El enunciado *a*) es ante todo o falso o inofensivo. Es falso si pretende que los juicios morales *enuncian* las actitudes de quien los profiere en el sentido de enun-

ciar que quien los profiere tiene esas actitudes. Pues si así fuera, serían simplemente observaciones autobiográficas, reemplazables sin pérdida alguna por enunciados que sean explícitamente de la forma «Mi actitud hacia esto es...» o «Mi sentimiento ante esto...». Pero si así fuera, no existirían desacuerdos morales interpersonales; cuando dos personas expresan lo que normalmente consideraríamos que son puntos de vista en conflicto, no se trataría en absoluto de puntos de vista en conflicto, sino que todo sucedería más bien como si, yendo dos personas en un bote, una de ellas dijera que se siente enferma y la otra replicara, por su parte, que ella no. Pero es un hecho evidente que existen desacuerdos morales genuinos, y que los puntos de vista morales pueden entrar en conflicto. Los juicios morales (por lo menos a este respecto) tienen que significar lo que suponemos que significan; y lo que suponemos que significan, la forma en que los usamos, es tal que sus pretensiones no son meramente autobiográficas, sino una clase de pretensión que es rechazada por quien profiere un juicio moral contrario. Por tanto, no se limitan a describir la actitud que tiene el hablante.

Con este argumento, sin embargo, no se renuncia a la posibilidad de que las preferencias morales, pese a no ser directamente descriptivas, puedan, no obstante, tener la función de expresar la actitud del que las hace. Esta pretensión es en sí misma incontestable e inofensiva; pues en sí misma no tiene otro alcance que el de decir que de un hombre que hace un juicio moral puede afirmarse (si ese juicio es sincero) que está expresando su actitud en relación con una determinada cuestión moral, y esto no es sostener ningún punto de vista distintivamente subjetivista sobre los juicios morales: de un hombre que hace sinceramente un juicio de hecho puede decirse que está expresando su creencia sobre un asunto de hecho, y esto no es defender una concepción subjetivista de los juicios de hecho. La interpretación subjetivista entra en liza cuando se afirma que uno que hace un juicio moral está expresando su actitud, y que *esto es todo lo que hay que decir sobre el asunto*. En particular, la fuerza subje-

tivista de *a*) radica en que sugiere que en relación con las actitudes expresadas en los juicios morales no se plantea el problema de su rectitud o de su error, mientras que en relación con las creencias expresadas en los juicios factuales sí se plantea la cuestión de su verdad o su falsedad.

Ahora que la tesis ha vuelto a obtener una forma distintivamente subjetivista, hay que decir que es falsa. Pues —apelando una vez más a las formas en que los juicios morales se hacen y se tratan en la realidad— no es verdad que en relación con las actitudes morales no se plantee la cuestión de su rectitud o de su error. Una de sus notas distintivas, por oposición a las meras expresiones de gusto o preferencia, por ejemplo, en materia de alimentos, es que tomamos muy en serio la idea de que un hombre esté equivocado en sus puntos de vista morales; por supuesto que el mismo concepto de concepciones o *ideas* morales marca aquí una diferencia al inclinarse, como lo hace, más bien en la dirección de la creencia que en la dirección del mero gusto o preferencia. Es precisamente una nota de la moralidad el que *de gustibus non est disputandum* no es una máxima que se aplique a ella.

Contra esto podría replicarse que el hecho de que las actitudes morales puedan ser calificadas de «correctas» o «erróneas», y de que la cuestión de su «corrección» o de su «error» sea tomada en serio, no contribuye en ningún sentido definitivo a trascender el subjetivismo. No demuestra que las actitudes morales sean más que (meras) actitudes, sino que son actitudes que nos causan preocupaciones; que nos interesa asegurar una similitud de actitudes dentro de la sociedad. El uso del lenguaje de lo «correcto» y lo «erróneo» puede ser considerado como una parte del aparato cuya función es asegurar el acuerdo, discriminando a los que disienten, etc.; queda, pues, en pie el hecho de que con lo único con que contamos es con actitudes de la gente hacia diferentes clases de conducta, personalidad, instituciones sociales, etcétera.

Parece, empero, que esta explicación no se ajusta a

los hechos: tiene por lo menos que subestimar la complejidad de la situación. En particular no logra dar razón del hecho indudable de que un hombre puede estar en un estado de duda moral, duda que puede resolver, de que un hombre puede cambiar de forma no arbitraria de ideas en relación con un asunto moral, y no meramente en el caso individual, sino también en relación con un asunto general, y todo ello por razones. Así, un hombre que antes estuviera convencido de que la actitud permisiva en relación con el aborto era errónea, puede cambiar de ideas en relación con ese asunto, y no meramente (por ejemplo) porque se sienta solo en un grupo que mantiene una actitud permisiva. No cabe duda de que muchos de los autores que escriben sobre filosofía moral sobreestiman la medida en que la gente se ve inducida a modificar sus ideas morales por consideraciones racionales; estos autores parecen ignorar la evidencia de hasta qué punto las actitudes son modificadas por factores tales como el deseo de conformarse con las ideas de un grupo o disentir de las de otro —hablamos, claro está, de los casos en que la pertenencia a un grupo no se decide por razones morales, sino que más bien viene determinada por la situación y las necesidades de los individuos. Pero esto, en un cierto sentido, cae fuera de nuestro asunto. Pues aun en el caso en que las actitudes morales rara vez estuvieran *determinadas* por razones, y las razones dadas en su apoyo sólo fueran racionalizaciones, nuestro modelo de las actitudes y de los juicios morales tiene que ser por lo menos suficientemente complejo para dejar sitio a las racionalizaciones. Pues, solamente si la postura a que se ve llevado un hombre por esas fuerzas satisface algunas condiciones que permiten considerarla como la clase de posición para la que las razones son relevantes, solamente entonces podremos entenderla como una postura moral.

Sin embargo, aun cuando las actitudes morales tengan la característica de admitir apoyo o ataque por medio de razones, y un despliegue de razones que termina en una conclusión; aun en ese caso, cabe todavía decir que estas actividades sólo son posibles dentro de un marco de

suposiciones dadas. Si las personas pueden discutir entre sí sobre una cuestión moral particular o sobre una cuestión de principio es porque existen actitudes morales de fondo que no se ponen en cuestión, y a la luz de las cuales es posible la argumentación. Dando a este punto una forma más fuerte, cabría decir que la argumentación sólo es posible porque se trata de la *aplicación* de ideas morales aceptadas. Y así, cuando no hay un trasfondo de acuerdo moral, no puede haber argumentación alguna. En este punto puede reaparecer la actitud subjetivista, ahora con la pretensión de que todo lo que se ha demostrado con estas consideraciones relativas al intercambio de razones es que la moralidad de un hombre o de una sociedad es hasta cierto punto general y sistemática y que las actitudes generales pueden ser aplicadas a casos menos generales. Cuando nos salimos de esa trama de actitudes generales en las que se está de acuerdo, acaba también la argumentación, y ya no hay forma de mostrar que una determinada postura es correcta o errónea.

Esto está muy cerca de nuestra segunda formulación, *b*), del subjetivismo; nos hemos visto conducidos a ella, según parece, introduciendo modificaciones en la primera. Sería deseable, no obstante, cambiar un elemento que parecía encajar de forma natural en la formulación de *b*), y con ello, cambiar también una implicación del término «subjetivismo». Pues, cuando se aborda el asunto de los desacuerdos últimos, es natural considerar como unidad que mantiene un conjunto de actitudes morales a la sociedad más bien que al individuo —no para hipostasiar a las sociedades, sino para llamar la atención sobre el hecho de que hay límites para el grado de desacuerdo último que puede existir dentro de una sociedad (pues sin cierto grado de homogeneidad moral no sería una sociedad); pero no hay límites, por lo menos de esta clase, para el desacuerdo entre sociedades.

Interludio: relativismo

Tomémonos un breve descanso en este punto de nuestra argumentación sobre el subjetivismo y echemos una ojeada a una idea o ensamblaje de ideas que se ha ido montando en el terreno de los desacuerdos morales entre sociedades. Se trata del *relativismo*, la herejía de los antropólogos, posiblemente la idea más absurda que se haya defendido jamás en filosofía moral. En su forma vulgar e irredenta (que es la que voy a considerar, ya que es la forma más señalada y también la más influyente) consiste en tres proposiciones: que «correcto» significa (sólo puede ser coherentemente entendido si significa) «correcto para una sociedad dada»; que «correcto para una sociedad dada» ha de ser entendido en sentido funcionalista; y que (por consiguiente) es incorrecto que los individuos de una sociedad condenen, se interfieran en, etc., los valores de otra sociedad. Idea con larga historia, fue popular entre algunos colonialistas liberales, sobre todo administradores británicos en lugares (tales como el África occidental) en que los blancos no tenían tierras. En este papel histórico, puede haber tenido, al igual que otras doctrinas confusas, una influencia beneficiosa, aunque es posible que el moderno nacionalismo africano pueda muy bien deplorar sus implicaciones conservadoras y tribalistas.

Cualesquiera que sean sus resultados, la idea es claramente inconsistente, ya que en su tercera proposición hace una afirmación sobre lo que es correcto y erróneo en el trato de uno con otras sociedades, y para ello se

hace uso de un sentido *no-relativo* de «correcto» no permitido por la primera proposición. La pretensión de que los sacrificios humanos, por ejemplo, eran «correctos para» los Ashanti, se acaba tomando en el sentido de que los sacrificios humanos eran correctos entre los Ashanti y esto, a su vez, en el sentido de que los sacrificios humanos entre los Ashanti eran correctos; esto es, que no es de nuestra incumbencia interferirnos en ellos. Pero esto último no es ciertamente la clase de pretensiones que la teoría permite. A lo más que la teoría puede dar cabida es a la pretensión de que es correcto para (es decir, es funcionalmente valioso para) nuestra sociedad la no interferencia con la sociedad de los Ashanti, y, primero, esto no es, ciertamente, todo lo que se estaba diciendo, y, segundo, es muy dudoso que tal cosa sea verdad.

Aparte de la vinculación, lógicamente desafortunada, que este punto de vista establece entre una moralidad no-relativa de la tolerancia y la no interferencia y una concepción de la moralidad como algo relativo, la teoría adolece en sus aspectos funcionalistas de alguna notoria debilidad del funcionalismo en general, en concreto, de las dificultades con que implica la identificación de «una sociedad». Si se considera a la «sociedad» como una unidad cultural, identificada en parte por sus valores, entonces muchas proposiciones funcionalistas dejan de ser proposiciones empíricas y se convierten en puras tautologías: por supuesto, y con ello nada se dice, que es condición necesaria para la supervivencia de un grupo-con-ciertos-valores que el grupo mantenga esos valores. En el extremo opuesto, la supervivencia de una sociedad podría ser considerada como la supervivencia de ciertas personas y el que esas personas tengan descendientes, y en este caso muchas de las proposiciones funcionalistas sobre la necesidad de la supervivencia cultural serán falsas. Cuando en Gran Bretaña algunos nacionalistas galeses hablan de la supervivencia de la lengua galesa como una condición de la supervivencia de la sociedad galesa, tratan a veces de trasmitir la impresión de que se trata de una condición de supervivencia de

las gentes de Gales, como si el olvido del galés fuera literalmente letal.

Entre estos dos extremos se encuentra el territorio realmente interesante, una provincia de fructífera ciencia social, en la que hay espacio para pretensiones tales como la de que una determinada práctica o una determinada creencia está integralmente conectada con muchos más aspectos de la fábrica de una sociedad que lo que podría parecer a primera vista, que no se trata de una excrecencia, de modo que el estorbarlas o el modificarlas llevaría a cambios sociales de alcance mucho más amplio que el que cabría haber esperado a primera vista; o también, que un cierto conjunto de valores e instituciones puede ser tal que, si se perdieran o cambiaran profundamente, la gente en esa sociedad, aunque sobreviviera físicamente, lo haría solamente en condiciones de desenraizamiento y desesperación. Tales proposiciones, si se confirman, serían de primera importancia a la hora de decidir qué hay que hacer; pero no pueden hacerse cargo de la tarea de decidir qué es lo que hay que hacer.

Aquí, y a todo lo largo de las cuestiones relativas al conflicto de valores entre sociedades, necesitamos (aunque rara vez se consigue) alguna imagen moderadamente realista de qué decisiones podrían ser tomadas y por quién, y de las situaciones para las que estas consideraciones podrían ser relevantes en la práctica. De los distintos paradigmas que vienen a la mente, uno es el del conflicto, tal como la confrontación de otras sociedades con la Alemania nazi. Otro es el del control, en el que (para evitar otras complicaciones del caso más obvio, el colonialismo) podríamos tomar el caso de las relaciones del gobierno central de Ghana con los elementos residuales de la sociedad tradicional de los Ashanti. En ninguno de los dos casos las proposiciones funcionalistas proporcionarían por sí solas ninguna respuesta en absoluto. Y mucho menos lo harán cuando la cuestión principal es la de si es realista o deseable considerar a un grupo dado como «una sociedad» en un sentido relevante, o de si sus valores y su futuro ha de considerarse integralmente referido a los de un grupo más amplio

—como es el caso de los negros en los Estados Unidos.

La confusión central del relativismo es la de tratar de extraer, a partir del hecho de que las sociedades tienen diferentes actitudes y valores, un principio *a priori* no-relativo para determinar la actitud de una sociedad hacia otra; esto es imposible. Si vamos a decir que existen desacuerdos morales últimos entre sociedades, hay que incluir, entre los asuntos en que tales desacuerdos pueden producirse, sus propias actitudes hacia las concepciones morales de los otros. También es verdad, sin embargo, que existen rasgos inherentes a la moralidad que tienden a hacer difícil la idea de que la moralidad se aplica solamente a un grupo. El elemento de la universalización que está presente en toda moralidad, pero que bajo la moralidad tribal se aplica solamente, quizás, a los miembros de la tribu, llega a hacerse progresivamente extensivo a todas las personas en cuanto tales. Dicho de manera menos formal: como ya hemos subrayado antes, es esencial a la moralidad y a su papel en toda sociedad que ciertas clases de reacciones y motivaciones queden fuertemente internalizadas, y éstas no pueden evaporarse sin más por el hecho de que uno se vea confrontado con seres humanos en otra sociedad. Al igual que *de gustibus non est disputandum* no es una máxima que se aplique a la moralidad, tampoco lo es la de que «cuando estés en Roma, haz lo que los romanos», que, como mucho, no es más que un principio de etiqueta.

Exactamente, no es cuestión de hacer lo que los romanos hacen, sino de si se soporta o no lo que hacen. Aquí sería una perogrullada subrayar que, por supuesto, alguien que llega a tener una experiencia más amplia del mundo puede, con razón, llegar a considerar como provincianas algunas reacciones morales suyas ante formas de conducta que no le resultan familiares, y tratará de modificarlas o de desconfiar de ellas. Existen muchas distinciones importantes que hacer aquí entre las clases de pensamientos apropiadas a tal proceso en los diferentes casos: a veces podrá ser que ya no considere en absoluto cierto asunto como una materia moral, a ve-

ces puede que llegue a la conclusión de que lo que fuera le pareció lo mismo que algo que hubiera deplorado en su propio país, era en realidad, en aspectos moralmente relevantes, una cosa muy diferente. (Quizás —aunque cuesta trabajo creerlo— hubo misioneros y otros que a los hombres en una sociedad con poligamia los consideraron bajo el prisma de bígamos en germen en el propio país.) Pero sería una concepción muy particular de la moral, e implausible tanto desde un punto de vista psicológico como desde un punto de vista moral, el insistir en que estas reacciones de adaptación eran las únicas correctas, el insistir en que, en la confrontación con prácticas que se creen inhumanas y como tales se perciben, hay una exigencia *a priori* de aceptación. En el fascinante libro de Bernal Díaz del Castillo, que estuvo con Cortés en Méjico, se narra lo que todos sentían cuando llegaban a los templos donde se practicaban sacrificios humanos. Este grupo de valientes, muy lejos por cierto de considerarse perfectos desde un punto de vista moral, se sentían genuinamente horrorizados por las prácticas de los aztecas. Sería sin duda absurdo considerar esa reacción como algo meramente provinciano o santurrón. Esa reacción indicaba algo que su conducta no indicaba, que consideraban a los indios como hombres y no como animales salvajes.

No es deshonesto insistir en esta clase de casos y, en general, en los casos de confrontaciones que hayan tenido lugar en la realidad. «Cada sociedad tiene sus propios estándares» puede ser, aunque vaga y confusa, una máxima útil a veces para el estudio de la sociedad; como máxima para el estudio de la sociedad, es también indolora. Pero después de todo, ¿qué se supone que va a hacer uno cuando se ve confrontado con un sacrificio humano? —una cuestión que no tiene visos de hacerse tangible para la mayoría de nosotros, quizás, pero bien real para Cortés. «No era asunto suyo», podría replicarse, «y en todo caso, no tenían ningún derecho a estar allí». Tal vez —aunque esto, una vez más, es necesariamente en sí mismo un juicio moral no-relativo. Pero aun cuando no tuvieran ningún derecho a estar allí, merece la

pena abordar en términos de una discusión moral efectiva la cuestión de qué se *seguiría* de ello. Si un ladrón sorprende al dueño de la casa en el momento en que trata de asesinar a alguien, ¿está moralmente obligado a no interferirse por el hecho de que su presencia allí sea indebida?

Nada de esto significa que queramos negar el hecho obvio de que muchos se han interferido con otras sociedades cuando deberían no haberlo hecho; se han interferido sin comprensión; y a menudo, con una brutalidad mucho mayor que la de nada de lo que trataban de atajar. Lo único que estoy tratando de decir es que no puede considerarse como una consecuencia de la naturaleza de la moralidad misma el que ninguna sociedad deba interferirse nunca en los asuntos de otra, o que los individuos de una sociedad, cuando se ven confrontados con las prácticas de otra, si son racionales, deban reaccionar con la aceptación. Sacar estas consecuencias es el paso característico (e inconsistente) que da el relativismo vulgar.

Subjetivismo: otras consideraciones

Los cebos del relativismo vulgar no son muy tentadores. Pero su error central es muy importante; y el que se trate de un error tiene una significación que trasciende esta doctrina particular, preocupada como está por las relaciones entre sociedades. Es posible que alguien persuadido por los subjetivistas deje de preocuparse de los problemas morales. (Que es cosa muy diferente de dejar de considerar algo como problema realmente moral. Así, es tanto posible como razonable suponer que no exista una moral distintivamente sexual, en el sentido de consideraciones morales que gobiernen las relaciones sexuales, y nada más; admitiendo, sin embargo, el hecho extremadamente obvio de que las relaciones morales son profunda y especialmente susceptibles de dar lugar a problemas morales, de confianza, explotación, despreocupación por los intereses de terceras partes, etc.) Un hombre confrontado con una injusticia política monstruosa, por ejemplo, puede que desconfíe de protestar o de luchar contra ella, pues, dirá, «¿Quién puede juzgar?» o «Se trata sólo de mis sentimientos frente a los de ellos», o algo parecido.

En la medida en que exista un nexo intelectual discernible entre el subjetivismo y la indiferencia, tiene que implicar algo parecido al error relativista. Pues la idea tiene que ser algo así como «Puesto que el subjetivismo es verdadero, no está justificado que yo proteste». Si esto es correcto, tiene que serlo, o bien porque si el subjetivismo es verdadero, nadie está justificado a hacer nada,

o porque, si el subjetivismo es verdadero, nuestro hombre está particularmente no justificado al protestar. Si se trata de lo primero, entonces los que cometen la injusticia tampoco están justificados en lo que están haciendo, ni tampoco tendrá justificación nuestro hombre para *no* protestar, y estas consideraciones quitan toda la *base* que se suponía tenía nuestro hombre para su indiferencia. En cualquiera de los casos la argumentación en esta versión ha ido sin duda demasiado lejos, incluso para el subjetivismo; pues el subjetivismo no pretendía que fuera imposible pensar que nada estuviera justificado, sino solamente que cuando un hombre dado piensa que algo está justificado, en última instancia no se le puede demostrar que está equivocado. El hombre con el que estamos discutiendo, en cierto modo, ha pasado de eso a una posición en la que nadie (ni siquiera subjetivamente) puede pensar que algo esté justificado, y eso tiene que ser un malentendido; a menos, por supuesto, que el subjetivismo sea inconsistente, en caso de lo cual el subjetivismo es falso, y entonces la discusión, de cualquier modo, ha terminado.

Pasemos, entonces, a la segunda alternativa, la de que, puesto que el subjetivismo es verdadero, entonces nuestro hombre está particularmente no justificado al protestar. ¿Por qué habría de ser esto así? «Bien», podría decirnos, «ellos piensan que tienen razón, y ¿quién soy yo para decir que no la tienen?». Pero la fuerza aparente de esta respuesta deriva toda ella del sutil abandono de la arena subjetivista y de la introducción de la idea de que existe algo así como una rectitud objetiva, solamente que no se está seguro de si las acciones de las otras personas la poseen o no. Si se empecina en la senda del subjetivismo, nuestro hombre tiene que reconocer que si decide pensar que los otros están equivocados y que tiene razón al protestar, entonces tampoco puede decir nadie que él está equivocado, con lo que él no está menos justificado al protestar que lo que lo están los otros al hacer las cosas que hacen. Otra forma de expresar este punto es ésta: «tal vez ellos tengan razón» tiene que ser uno de *sus* pensamientos morales

(de nuestro hombre, se entiende). Si también tiene el pensamiento «Están equivocados (sólo que yo no estoy justificado al protestar)», nuestro hombre tiene pensamientos morales inconsistentes dentro de su propio sistema, y el subjetivismo nunca le exigió que hiciera esto. Pues no cabe duda de que el pensamiento «¿Quién soy yo para decir que están equivocados?» es un pensamiento que ha tenido, no dentro de su propio recinto subjetivo, sino en el espacio intermedio entre su propio recinto y el de otras personas; es un pensamiento que trata de mantenerse fuera de todas las posiciones morales (incluyendo la del que lo tiene) y seguir siendo, no obstante, un pensamiento moral. Pero este lugar intermedio, aun para el propio subjetivismo, no es lugar en el que nadie pueda tener un pensamiento moral.

Esta clase de razones muestran por qué los empeñados en desactivar el subjetivismo dicen que el subjetivismo deja todo donde estaba y que no es posible que lógicamente pueda implicar indiferencia ni ninguna otra actitud práctica. (Quizás con demasiada rapidez, pasan de ahí a la pretensión de que sus ideas no pueden *alentar* ninguna de esas actitudes, basándose tal vez en el cuestionable argumento de que a mí no se me puede hacer responsable de las confusiones en que cae la gente como resultado de lo que digo.) Pero, ¿el subjetivismo deja realmente todo donde estaba? Ciertamente que no *todo*. Las marañas dialécticas que acabamos de atravesar son, a mi juicio, perfectamente válidas en el sentido de que muestran que la indiferencia no podría seguirse lógicamente del subjetivismo. Pero la razón por la que esto es así es que la argumentación en favor de la indiferencia exige, y el subjetivismo prohíbe, lo que hemos denominado la posición intermedia. Pero al prohibir esa posición intermedia, el subjetivismo *parece* haber quitado algo, pues por lo menos parece que somos capaces de reconocer esa posición intermedia (por ejemplo, en los mismos enunciados del subjetivismo). Otra forma de expresar nuestro malestar es que parece que en estos argumentos no se nos ha dado ninguna razón especial de por qué la posición intermedia está vetada a la *moralidad*.

Dijimos que la posición intermedia no era sitio donde pudiera tener lugar un pensamiento moral. Pero sí que parece un sitio apropiado para otras clases de pensamiento, en particular, para el pensamiento factual. Por supuesto que el subjetivismo (véanse *b*) y *c*) de las formulaciones iniciales que dimos anteriormente) insiste en el contraste entre las actitudes morales, considerando las últimas como «objetivas» en un sentido en que no lo son las primeras. El subjetivista nos deja, y puede que pretenda dejarnos, con la molesta sensación de que las creencias factuales tienen, y a las morales les falta, algo que es bonito tener: con la molesta sensación de que las creencias factuales y la ciencia son en cierto modo *más sólidas* que la moralidad.

Pero aun así, se nos dirá, el subjetivismo deja todo donde estaba en lo que concierne a la moralidad, aunque tal vez no en lo que concierne a los sentimientos confusos sobre la moralidad. Pues aun supuesta la diferencia de solidez, de ello no puede sacarse ningún género de conclusión práctica. En particular, no podemos concluir (como parece obvio que algunos se sienten inclinados a hacer hoy) que, puesto que la ciencia es objetiva y la moralidad no lo es, estamos objetivamente justificados a *dedicarnos* a la ciencia, mientras que sólo lo estamos subjetivamente para protestar contra la injusticia. Pues dedicarse a la ciencia no es una actividad menos práctica que cualquier otra, y no hay más razón de por qué una habría de estar objetivamente justificada y no cualquier otra. Las justificaciones para dedicarse a asuntos objetivos no son justificaciones objetivas para dedicarse a esos asuntos, como tampoco el hecho de que existan justificaciones deductivas de los teoremas de *Principia Mathematica* significa que existan justificaciones deductivas de los proyectos de leer, reconstruir, o descubrir los teoremas de *Principia Mathematica*. Todos estos son ejemplos de esa falacia particularmente tentadora a la que podemos llamar el principio de los «bueyes gordos»: para conducir bueyes gordos, tiene uno que ser gordo.

De la misma manera, aunque de forma más sutil, el hecho de que las medidas políticas «pragmáticas» (esto

es, medidas políticas en las que se hace uso de sofisticadas consideraciones técnicas en la persecución del propio interés) puedan ser comparativamente evaluadas entre sí con mayores expectativas de que los expertos lleguen a un acuerdo que lo que cabe esperar en el caso de medidas más idealistas, no demuestra, aunque sea verdadero, que tengamos razones más objetivas y sólidas para adoptar políticas pragmáticas que para adoptar políticas idealistas. Una vez que las hemos adoptado, puede que tengamos razones más objetivas a la hora de desarrollarlas, o sea, a la hora de decidir qué es lo que hay que hacer a continuación; pero nunca pisamos sobre razones más objetivas con referencia a la cuestión de si la clase de cosas que estamos haciendo son las que debemos hacer o no. Simplemente nos sentimos más cómodos al trabajar, y tanto más si lo hacemos entre expertos sofisticados.

Así el subjetivismo, incluso con este molesto contraste suyo entre ciencia y moralidad, lógicamente aun no tiene por qué extraer de ello consecuencias prácticas. Ni incluso puede arrojar la conclusión de que estamos más sólidamente justificados para *tener* creencias factuales que creencias morales; o de que pisamos sobre razones más objetivas cuando se trata de una cuestión factual que cuando se trata de una cuestión moral; o de que estamos objetivamente justificados para buscar la verdad sobre algo; o para tratar de buscar explicaciones científicas de los fenómenos y no contentarnos con la superstición; pues todo lo que el subjetivismo nos dijo es que las creencias factuales y científicas eran objetivas; ahora bien, el que hayamos de buscar creencias factuales o científicas no es, a su vez, una creencia factual o científica.

Tal vez el subjetivista conceda de grado estas observaciones relativas a la dedicación a la actividad científica o al pensamiento factual racional. Examinemos nosotros, pues, con más detalle esa diferenciación central que establece entre esas actividades por un lado, y el pensamiento moral por el otro. El subjetivista puede muy bien empezar diciendo que si nos dedicamos a esas actividades, entonces, como los hechos son lo que son,

no tenemos más remedio que llegar a alguna clase de acuerdo sobre las creencias científicas o factuales; ahora bien, porque nos introduzcamos en una discusión moral, y los hechos sean lo que son, no por eso puede decirse que no tengamos más remedio que llegar a algún acuerdo en las creencias morales. Existe un elemento en esta respuesta que debe corregirse de una vez por todas. Pues no es en modo alguno verdad que, porque los hechos sean como son, no tengamos más remedio (supuesto el pensamiento factual) que llegar a algún acuerdo en las creencias factuales: los hechos pueden estarnos ocultos. Lo más que podemos decir es que si reconocemos que los hechos son así, no tenemos más remedio que llegar a algún acuerdo sobre las creencias factuales. Y esto es una tautología, pues el reconocimiento por nuestra parte de que los hechos son así implica el que lleguemos a algún acuerdo sobre las creencias factuales. Tal vez el subjetivista pueda corregir el escaso interés de esta proposición diciendo algo parecido a lo que sigue: que si dos observadores están en la misma situación de observación, y tienen los mismos conceptos y no son deficientes como observadores, etc., entonces llegarán a las mismas creencias factuales sobre esa situación. Si el «etc.» en esto último puede ser manejado de forma que la proposición resulte verdadera, entonces es casi seguro que acabará siendo también necesariamente verdadera. Muy bien, responde el subjetivista; pero *eso* no es necesariamente verdadero, y por supuesto, no es verdadero en absoluto que si dos observadores están en la misma situación de observación, tienen los mismos conceptos, etcétera, entonces no tienen más remedio que llegar a las mismas creencias *morales* —y de ahí el contraste. Pero, puede ser nuestra réplica, ¿estamos seguros de que esto último no es verdad si suponemos que tienen los mismos *conceptos morales* —que es sin duda la verdadera analogía? No, no es verdad ni siquiera en ese caso, dice el subjetivista; pues si tienen, como muy bien pudiera ser, unos conceptos morales más bien mínimos, tal como el mero concepto de *lo que uno debe hacer*, pueden estar de acuerdo en todos los hechos y, sin embargo, disentir moralmente.

Esta es una tesis central de muchos filósofos que insisten en la distinción hecho/valor. Estos filósofos se representan las diferentes concepciones morales de la siguiente forma: todas ellas harían uso de algún concepto moral común, a modo de esqueleto, que rellenarían de forma diferente o al que darían diferentes contenidos. Ahora bien, cabría insistir en que buena parte del pensamiento moral no opera con esos conceptos empobrecidos; en que, por lo común, pensamos en términos de ideas más complejas de virtudes, tipos de acción equivocada, etc., nociones, por ejemplo, de *robo*, *cobardía*, o de *lealtad*, o de los *deberes de la profesión de uno*. Con estos conceptos morales más sustanciales, existen mejores expectativas de que, si usamos unos mismos conceptos, llegaremos a un acuerdo, o por lo menos, a los tipos de desacuerdo que razonablemente podemos esperar, y que son menos desalentadores. Pero el subjetivista dirá que sigue existiendo la asimetría siguiente: que en los contextos científicos o factuales, si dos observadores tienen diferentes conceptos pueden, en última instancia, o bien darse cuenta de que los conceptos son de hecho equivalentes, o si no, encontrar, para preferir los unos a los otros, una razón en términos de éxito predictivo, potencia explicativa, etcétera, y que en el caso de la moralidad no existe nada paralelo a esto.

Aun cuando abandonemos la ingenua creencia (a la que algunos subjetivistas se aferran cuando insisten en este contraste) de que la ciencia «prueba» cosas; aun cuando aceptemos que lo que la ciencia hace es eliminar hipótesis y que existen infinitas hipótesis que nunca han sido eliminadas porque son demasiado excéntricas como para que nadie se tome la molestia de comprobarlas (una observación que he oído hacer a Hilary Putnam); sin embargo, existen procedimientos de eliminación que tienen que ser respetados por las personas que se mueven en el contexto de la ciencia, e incluso los desacuerdos conceptuales pueden, con suerte, quedar sometidos a criterios que son aceptados de forma impersonal. Pero para algunos desacuerdos entre personas y grupos, cuando ambos se ven envueltos en discusiones que se

admiten como morales, no existen tales procedimientos. El contraste —algún contraste— existe.

Pero, ¿por qué no habría de existir? Este es el punto en el que finalmente el filósofo que durante todo este tiempo ha estado discutiendo con el subjetivista podría por fin dar un giro y decir: por supuesto que el contraste existe; la moralidad no es ciencia o conocimiento factual, y es esencial que no lo sea. Lo propio de la moralidad no es reflejar el mundo, sino cambiarlo; la moralidad se refiere a cosas tales como principios de acción, elección, responsabilidad. El hecho de que hombres de igual inteligencia, conocimiento factual, etcétera, puedan disentir moralmente, nos dice algo sobre la moralidad: que (dicho aproximadamente) no podemos cargar la responsabilidad de lo moral sobre cómo es el mundo. Pero esto no demuestra (como el subjetivismo parecía en un principio insinuar) que, por ello, algo ande *mal* en el asunto.

Alguna afirmación de este tipo —y lo único que he hecho es esbozarla a grandes rasgos— tendrá que constituir la culminación del proyecto que he llamado de desactivación del subjetivismo. Y según lo veo yo, ha de concedérsele algún éxito. Las formas más obvias en las que alguien podría sentirse desalentado por el subjetivismo parecen haber quedado bloqueadas. Razón por la que, para volver una vez más sobre la cuestión del indiferentismo, que hemos tratado antes, el empeñado en este proyecto de desactivación podría reordenar alguno de los materiales del subjetivista en una argumentación como la siguiente: Observamos que, cuando hombres de igual competencia científica o histórica, de igual habilidad perceptiva o intelectual, etcétera, están en profundo desacuerdo sobre algún asunto científico o histórico, existe una buena razón para que dejen de disentir de forma tan profunda y para reconocer algo que su mismo desacuerdo, supuestos sus conocimientos y su capacidad, pone de manifiesto, a saber: que el asunto es *incierto:* entonces, es racional tanto para ellos como para terceras partes el suspender el juicio. Uno se siente tentado de pensar que lo mismo cabría decir de los casos

de desacuerdo moral; pero sería una equivocación. Ya que sería empezar estableciendo una diferencia entre la moralidad y el conocimiento factual, para acabar asimilándolos después. Pues la diferencia vital es que el desacuerdo en la moralidad se extiende a qué es lo que habría que hacer, y esto implica, por ambas partes, el preocuparse por lo que ocurre; y una vez que uno se da cuenta de esta diferencia, se da cuenta igualmente de que no podría constituir una exigencia de la racionalidad el que uno dejara de preocuparse de estas cosas porque el otro no está de acuerdo con uno.

Este argumento, y otros similares, me parece que demuestran que la operación de desactivación se ha llevado a cabo con éxito en ciertos aspectos vitales. ¿Ha sido completo el éxito? Si lo ha sido, entonces estábamos equivocados cuando hace un rato nos sentíamos incómodos porque el subjetivismo excluía para la moralidad, pero no para las creencias factuales, algo que llamábamos la «posición intermedia». Pero a mi juicio no estábamos del todo equivocados cuando dábamos pábulo a ese sentimiento de malestar —tal vez ahora podamos ver por qué.

Si tengo un desacuerdo factual con un hombre, podría pensar: «Yo creo que *p*, él cree que no-*p*. Quizás tenga razón.» La forma natural de entender esto es como expresión de duda, como expresión de una confianza en cierto modo disminuida; la forma en que me digo «Quizás tenga razón» implica que si bien yo sigo creyendo todavía que *p*, con todo, mi creencia no es ya tan fuerte. Si lo entendemos de esta forma, puede establecerse sin esfuerzo un paralelo con el caso moral; pues si pienso «Quizás tenga razón», esto será tomado también como expresión de una confianza en cierto modo disminuida. Ahora bien, en el caso factual es posible un pensamiento que parece similar pero que no es el mismo: el pensamiento «Yo estoy convencido de que *p*, pero, con todo, es posible que no-*p*», y esto no es en absoluto una expresión de *duda*, sino que más bien registra la consideración impersonal de que lo que las cosas son es independiente de mis creencias; ellas son lo que son,

cualesquiera sean mis creencias. No sabemos exactamente qué contenido asignar a este pensamiento, pero a menos que sostengamos los puntos de vista filosóficos más drásticos, estamos convencidos de que ese pensamiento tiene un contenido: el empeñado en la tarea de desactivación estará de acuerdo en esto. Pero, precisamente, el subjetivismo desactivado no permite pensamiento paralelo posible por el lado moral: para el subjetivismo, aunque desactivado, justamente no existe contenido para «Yo estoy convencido de que la discriminación racial es mala, pero, con todo, es posible que no lo sea», excepto cosas tales como «¿Hasta qué punto estoy convencido?» o «Supongo que alguien podría hacerme cambiar de parecer».

Tal contraste (y, lo mismo éste que otros contrastes relacionados con él, necesitan de una adecuada investigación que no podemos emprender aquí) podría hacer que estuviéramos de acuerdo con la tercera fórmula subjetivista que introdujimos al principio: no hay hechos morales. Sin embargo, una vez más, el empeñado en la desactivación dirá: esto no es sino una formulación más de lo que ya he dicho, y se trata de una formulación de algo esencial y no nocivo para la moralidad. Pues yo ya he dicho que el pensamiento moral es esencialmente práctico; y que no es asunto suyo reflejar el mundo. Pero ahora podríamos replicar: dijiste que no era asunto suyo reflejar el mundo de los hechos empíricos y estábamos de acuerdo en ello. Pero, ¿estábamos de acuerdo acaso en que no reflejaba hechos en absoluto? Y aquí el punto que nos dejaba insatisfechos adquiere más claridad con la idea de que la razón por la que incluso el subjetivismo desactivado parece haber dejado algo fuera es que el pensamiento moral *parece como si* reflejara algo, como si estuviera constreñido a seguir algo más bien que a ser libremente creador. Cuando vemos, además, que muchos filósofos empeñados en esta tarea de desactivación expresan la esencial diferencia entre el pensamiento moral y el pensamiento factual en términos de un contraste entre el entendimiento y la *voluntad*, y se representan el terreno de la moralidad en términos

de una *opción* por determinados principios morales, entonces tenemos razones para sentirnos insatisfechos, o bien con ellos, o bien con el pensamiento moral. Pues ciertamente, la conciencia de un principio de acción por el que libremente se opta es difícil que sea la conciencia de un principio moral, que es más bien la conciencia de algo que tiene que ser *reconocido*. Si además se nos dice que eso tiene su explicación psicológica, entonces el pensamiento moral parece un engaño, al presentársenos como algo muy parecido a algo que no es.

Estas observaciones se limitan a apuntar hacia el centro de nuestra insatisfacción. Dejan casi todo por hacer: y tal vez no sólo en filosofía moral. Por ejemplo, necesitamos distinguir dos cosas que van mezcladas en lo que acabamos de decir, la idea del realismo —que el pensamiento tiene un objeto que es independiente del pensamiento— y la idea de que el pensamiento esté constreñido a sacar ciertas conclusiones. Así, el pensamiento matemático tiene esta última propiedad, pero es una cuestión profunda e incierta en la filosofía de la matemática la de hasta qué punto habríamos de o podríamos pensar en términos realistas sobre el objeto de las matemáticas.

Vamos a dejar aquí la discusión directa del subjetivismo y de los temas que suscita, con la conclusión de que el subjetivismo desactivado no deja todo donde estaba, pero sí deja donde estaba más de lo que hubiera cabido pensar cuando iniciamos esta discusión. Si, aunque desactivado, hay subjetivismo, las cosas en la moralidad no son exactamente lo que parecían ser; pero el fraude, habría que añadir, justifica a lo más el malhumor, más bien que el pánico. No perderemos de vista, empero, la idea de ligaduras sobre el pensamiento moral, las limitaciones en la creación de valores. Volveremos a ellas por otro camino, a través de la idea de la bondad. Para hablar de la bondad, empezaremos hablando de «bueno».

«Bueno»

El uso del término «bueno» ha constituido un foco para muchas discusiones de asuntos básicos en filosofía moral; aunque sería un error pensar que esta palabra, o sus equivalentes aproximados en otras lenguas, podrían soportar ellos solos todo el peso de esa temática, sí es cierto que el examen de ese término representa una útil palanca para suscitar y clarificar algunos de ellos. Empezaremos con algunas consideraciones lógicas: éstas nos conducirán a cosas de mayor sustancia moral.

Como observó Aristóteles, «bueno» se dice de muchas clases de cosas; de cosas que, por supuesto, pertenecen a categorías diferentes. Y aunque en cierto modo no queremos decir lo mismo cuando lo aplicamos a estas diferentes clases de cosas —y en este sentido, lo que hace que un general sea un buen general es diferente de lo que hace que un médico sea un buen médico—, no por ello podemos decir que el término sea ambiguo: no podríamos reformar el lenguaje y decir justamente lo que queremos decir ahora si nos pusiéramos a sustituir «bueno» por una expresión diferente cada vez que el término aparece.

Más que una teoría, lo que se ha tratado en tiempos recientes es de construir un modelo para mostrar que «bueno» es genuinamente no-ambiguo. Una de esas tentativas fue la de G. E. Moore, quien sostenía que la bondad era una simple propiedad indefinible como la amarillez, sólo que, a diferencia de la amarillez, era *no-natural*, es decir (aproximativamente), que no era la clase de

propiedad cuya presencia o ausencia pudiera establecerse por investigación empírica, si bien (de una forma que queda muy oscura en esta teoría) la observación de las características empíricas de una cosa era, sin duda, relevante para la aprehensión de la bondad.

Aparte de su evidente carácter misterioso y no explicativo, este planteamiento de Moore se enfrenta a una objeción lógica. Un importante rasgo de «bueno» es que, en muchos de los casos en que aparece, funciona como un adjetivo *atributivo* y no como un adjetivo *predicativo* (que es la forma en que P. T. Geach ha expresado esta distinción)[1]. «Amarillo», por ejemplo, es un adjetivo predicativo, pues un enunciado de la forma:

Es un pájaro amarillo

admite el análisis

Es un pájaro y es amarillo.

Además, de los dos enunciados siguientes:

Es un pájaro amarillo
Un pájaro es un animal

podemos inferir la conclusión

es un animal amarillo.

Pero el enunciado

Es un buen jugador de criquet

no puede analizarse como

Es un jugador de criquet y es bueno

ni de los enunciados

Es un buen jugador de criquet
Un jugador de criquet es un hombre

podemos inferir válidamente la conclusión

Es un hombre bueno.

[1] P. T. Geach, «Good and Evil», *Analysis*, vol. 17 (1956).

A un adjetivo que tiene esta última característica, la de estar lógicamente pegado al sustantivo al que califica, podemos llamarlo adjetivo atributivo; o, hablando con más precisión, a una forma de uso en la que el adjetivo va pegado de ese modo al sustantivo al que califica, podemos denominarla forma atributiva de uso del adjetivo. Ahora bien, en la explicación de Moore se sostenía que «bueno» se parecía a «amarillo» en el hecho de que, al igual que éste, representaba una cualidad simple, pero que se distinguía de «amarillo» en que la cualidad por la que estaba era una cualidad no natural; y, por más misterioso que eso suene, tiene que implicar, por lo menos, que la conducta lógica de «bueno» como adjetivo tiene que ser igual que la de «amarillo». Pero no lo es, y, por tanto, la explicación de Moore tiene que ser rechazada no porque no sea esclarecedora y sugerente, sino porque está radicalmente desencaminada.

Otro importante adjetivo atributivo es «auténtico» —una aserción de que algo es auténtico sólo puede ser entendida si podemos responder a la pregunta, «¿auténtico qué?». Esto puede quedar ilustrado por esa situación en el mundo del arte, en la que los coleccionistas pueden estar interesados en adquirir las obras de ciertos falsificadores, de suerte que algunas personas pueden estar interesadas en falsificar falsificaciones: así, puede plantearse la cuestión de si este cuadro es un auténtico Van Meegeren, pues todo el mundo sabe ya que en cualquier caso no se trata de un auténtico Vermeer.

La característica de la atributividad necesita, sin embargo, de una exploración más profunda si queremos llegar a una comprensión de «bueno». Podemos comprobar que «grande» es atributivo: así, de

a) Es un ratón grande

b) Un ratón es un animal

no podemos inferir válidamente

c) Es un animal grande.

La explicación del fallo de esta inferencia, y de la atri-

butividad de «grande» es bastante clara —se trata de que «grande» funciona como término comparativo, y «Es un ratón grande» significa algo así como «Es un ratón de mayor tamaño que la mayoría de los ratones». Si hacemos un análisis similar, la conclusión *c*) viene a significar algo así como «Es un animal de mayor tamaño que la mayoría de los animales», y entonces podemos ver por qué falla la inferencia. Lo que podría inferirse válidamente de esas premisas, si atendemos al análisis comparativo, es

d) Es un animal más grande que la mayoría de los ratones

que, como se ve, es una conclusión correcta.

¿Podemos explicar de la misma forma la atributividad de «bueno»? Es difícil que toda atributividad pueda explicarse en esos términos —un Van Meegeren auténtico no es más auténtico que la mayoría de los Van Meegeren. Parece más plausible sugerir que «bueno» o «buen» en «buen F» es atributivo porque «buen F» significa algo así como «mejor que la mayoría de los Fs». Sin embargo, un examen más atento demuestra que no es así. Acabamos de ver que basándonos en el análisis comparativo de «grande», podemos pasar válidamente de las premisas *a*) y *b*) a la conclusión *d*). Si «bueno» o «buen» fuera atributivo por ser comparativo, entonces, de forma similar, de

Es un buen jugador de criquet
Un jugador de criquet es un hombre

podríamos concluir:

Es un hombre mejor que la mayoría de los jugadores de criquet

pero esta conclusión no es válida, y está expuesta a las mismas objeciones que la inicial, «Es un hombre bueno». Estaríamos más cerca de una conclusión aceptable sólo con algo como lo siguiente:

Es un hombre mejor en el criquet *(o* mejor como ju-

gador de criquet) que la mayoría de los jugadores de criquet.

E incluso esto es dudoso, pues parece que no hay contradicción en la idea de que, con lo floreciente que es la situación de este juego, la mayor parte de los jugadores de criquet son bastante buenos; si esto es posible, el análisis comparativo en cualquier sentido parecido a la forma que acabamos de darle, desaparece también. Pero, aparte de esto y, en cualquier caso, el análisis comparativo no ha logrado despegar a «bueno» de su sustantivo; en la conclusión anterior, la conexión persiste, con «bueno» esta vez en su forma comparativa («mejor»). Así, la atributividad de «bueno» exige una conexión más íntima con su sustantivo que la exigida en el caso de un adjetivo meramente comparativo como «grande» («de gran tamaño»).

Como «bueno» en esta clase de construcción está íntimamente conectado con el sustantivo al que califica, el significado de una frase de la forma «un buen *x*» tiene que ser tomado como un todo; y su significado viene en parte determinado por lo que ocupe el lugar de *x*. ¿Podemos dar un paso más y decir que, en frases de esta forma, el significado del todo viene *esencialmente* determinado por el significado de lo que ocupe el lugar de *x*? En muchos casos parece como si pudiera darse este paso. Pues si consideramos definiciones funcionales de artefactos, tales como «reloj» o «abridor de latas», o también descripciones de seres humanos que se refieren a sus papeles, oficios o actividades especializadas, tales como «jardinero» o «general» o «jugador de criquet», parece que si uno entiende estas expresiones (por lo menos en el sentido fuerte de que uno entiende qué es un abridor de latas, por ejemplo, o qué hace un general), entonces uno comprende, dentro de ciertos límites, qué es una buena cosa de esta clase.

Puede tratarse de una comprensión en un nivel muy general y abstracto, y dentro de ese marco de referencia abstracto y lato quedará mucho espacio para el desacuerdo y para la sofisticada comparación de méritos. En particular, puede haber puntos de vista divergentes

o cambiantes en relación con qué aspectos de una actividad humana tienen que tener qué clase de peso en la evaluación: así, una persona, o una época, pueden tener puntos de vista distintos de los de otras en relación con la importancia que tiene para la estimación de un general como bueno el que gane sus victorias con un mínimo de pérdidas de vidas humanas. Ello no obstante, una comprensión de lo que es un *x* parece que ciertamente, en estos casos, contiene una comprensión general de los criterios apropiados para decir que algo es un buen *x*. Los casos más claros que todos son, por descontado, los de descripciones técnicas de artefactos; si alguien entrara en una factoría de aviones y dijera «esto es una buena ala sostenedora» refiriéndose a un prototipo que hubiera sido rechazado por estar mal diseñado o mal realizado, no cabe duda de que se habría equivocado; y si después explicara que dijo eso porque le encantó la forma o el pulimento del artefacto, eso en nada mejoraría la observación que había hecho, ya que éstos no son criterios para decidir si un ala es *una buena ala*, aunque muy bien pueden ser criterios apropiados para alguna otra evaluación de esa pieza de metal, por ejemplo, como objeto estético. (Esto ilustra una vez más la importancia que, con la evaluación, adquiere la cuestión de bajo qué concepto está siendo evaluada la cosa.)

Existe en la filosofía contemporánea una fuerte tradición de resistencia a la idea de que los criterios de valor, de lo que hace a una cosa de una cierta clase una buena cosa de esa clase, no pueden nunca ser lógicamente determinados por verdades factuales o conceptuales: esto es una aplicación central de la distinción entre hecho y valor a la que ya nos hemos referido. Esta resistencia fue influyentemente alentada por Moore, que acuñó la denominación de «falacia naturalista» para la equivocación en la que se pretendía que incurría cualquier punto de vista que sostuviera que la bondad de una cosa podía ser identificada con algún conjunto de características empíricas, o también metafísicas. Moore hizo una exposición de esta equivocación en términos de su teoría de que la bondad era una propiedad no

natural, punto de vista que, como ya hemos visto, en la medida en que en absoluto resulta comprensible, es un triste error. Muchos filósofos modernos que están de acuerdo en que Moore no tenía razón en su punto de vista, coinciden, sin embargo, con él en considerar que la «falacia naturalista» es una falacia genuina e importante; y dan una nueva explicación de su naturaleza. Esta explicación, reduciéndola a trazos demasiado sumarios, viene a ser la siguiente: que la función de los enunciados de la forma «esto es un buen *x*» es prescribir o recomendar, o la de llevar a cabo algún propósito lingüístico de este tipo en el marco general de lo normativo o evaluativo, mientras que el limitarse a describir las características de *x* no es llevar a cabo tal propósito; y ningún conjunto de enunciados que no lleven a cabo tal propósito puede lógicamente implicar ningún enunciado que lo haga. Prescribir, recomendar, etc., es hacer algo que los hechos por sí mismos no pueden hacer que hagamos; tenemos que tener alguna actitud evaluativa o prescriptiva que haga que resalten ciertas características, si es que esas características han de contar para nosotros como razones de aprobación. El mero conocimiento del mundo, o la comprensión de los conceptos, no pueden por sí mismos bastar para producir tal cosa.

Un examen completo de esta posición requeriria algo que nos llevaría muy lejos en este ensayo, el estudio de un importante campo en desarrollo en la filosofía del lenguaje, la teoría de los *actos de habla*, de las distintas cosas que podemos hacer al emitir oraciones. Sin embargo, podemos señalar brevemente tres puntos. El primero es que no puede existir ninguna conexión muy simple entre actos de habla tales como recomendar o prescribir por un lado, y el significado de enunciados tales como «esto es un buen *x*» por el otro. Como mucho, el uso de estos enunciados constituirá un acto de recomendación, etc., solamente si los enunciados son de hecho objeto de *aserción:* pero un enunciado tiene el mismo significado tanto si es objeto de aserción como si no lo es. Así, entendemos el enunciado «Es una buena película» en el contexto «Si es una buena película, ganará

algún Óscar»; pero en este contexto no es objeto de aserción y, por tanto, no ha tenido lugar ningún acto de recomendación[2]. Así pues, la conexión entre significado y recomendación tiene, cuando menos, que ser indirecta.

En segundo lugar, la teoría parece suponer demasiado deprisa que las funciones de recomendar, etc., y de describir, se excluyen entre sí. Sin embargo, una misma creación puede vehicular de una vez muchos actos de habla: si digo «Mañana va a llover», puedo, de una vez, haber descrito el tiempo de mañana, hecho un pronóstico, o hecho una advertencia a alguien, etc. Además, la cuestión de si efectivamente, al emitir una cierta oración, llevo a cabo alguno de estos actos de habla, puede venir determinada por lo que digo junto con los hechos de la situación. Así, si digo «El hielo es fino», sin duda alguna habré descrito el hielo, pero también, a la luz simplemente de los intereses y propósitos de alguien, puedo haber hecho algo que para ese alguien cuenta como una advertencia. De forma parecida (aunque no exactamente similar), si digo, describiendo este reloj que no atrasa ni adelanta, que no necesita cuerda, que nunca se estropea, etc., habré hecho algo muy parecido a recomendarlo como reloj, de forma completamente independiente de ninguna elección por mi parte de los criterios que deciden si un reloj es un buen reloj. Por descontado que las características de este reloj y la naturaleza general de los relojes no pueden hacer que yo pase a la recomendación —puedo permanecer en silencio—: pero determinan que *si* hago una descripción verdadera de este reloj en esos aspectos, entonces estoy haciendo algo en el ámbito de recomendar o de hacer una evaluación favorable de este reloj.

Este último punto nos lleva a la tercera consideración. Actividades tales como recomendar, etc., son ante todo actividades que se hacen abiertamente, conectadas con la efectiva emisión de oraciones: es por lo que he dicho que su estudio pertenece a la teoría de los actos

---

[2] Cfr. J. R. Searle, «Meaning and Speech-Acts», *Philosophical Review*, vol. 71 (1962).

de habla. Pero, posiblemente, ninguna explicación de los enunciados que contienen «bueno» pueda ser adecuada si se mantiene sólo en este nivel; pues cabe limitarse a pensar, o a creer, o a llegar a la conclusión de que algo es bueno en su clase, sin hacer en absoluto ninguna proferencia en ese sentido. Si aborrezco a Bloggs, puedo abstenerme cuidadosamente de hacer ninguna observación favorable de sus realizaciones como jugador de criquet, es decir, precisamente abstenerme de encomiarlo o de hacer nada parecido; sin embargo, en mi interior puedo verme forzado a reconocer que es un buen jugador de criquet. Una explicación adecuada tiene que dejar lugar para ese reconocimiento.

Estas consideraciones extremadamente sumarias pueden servir de indicación de algunas de las dificultades con que se enfrenta esta renovada versión de la proscripción de la «falacia naturalista». En realidad pienso que, como doctrina general de las funciones de «bueno», la distinción tajante entre hecho y valor que este punto de vista implica constituye un error: parece claro que para muchos argumentos de *x* en «esto es un buen *x*», una comprensión de lo que *x* es o hace, y el conocimiento factual de *x* —esto es, una combinación de información conceptual y factual— le bastan a uno para determinar, por lo menos a grandes rasgos, la verdad o la falsedad del juicio. Esto, ciertamente, es objetividad. Pero, ¿es así para todos los argumentos de *x*? Cuando examinamos esta cuestión, encontramos casos en los que todo resulta mucho más difícil, y en los que la distinción hecho-valor, o algo parecido, ofrece mejores asideros: casos en los que se hacen necesarias algunas consideraciones más profundas sobre el valor. Lo que buena parte de la filosofía contemporánea ha presentado como una doctrina lógica muy general sobre el comportamiento de «bueno» puede considerarse que tiene su interés como una doctrina mucho más restringida y no puramente lógica sobre la bondad de cosas y, más particularmente, de personas bajo ciertas descripciones específicas.

Bondad y roles

Consideremos la expresión «un buen padre»[3]. Aunque a grandes rasgos está medianamente claro qué clase de criterios acompañan a esta expresión, no está claro en absoluto que la comprensión de lo que es ser padre implique una comprensión de esos criterios. Y esto no es solamente porque la idea de ser un buen padre contenga una referencia a ciertas convenciones sociales; pues eso también ocurre con «jugador de criquet» o «empleado de banco» —saber lo que es un empleado de banco implica una buena dosis de conocimiento del tejido social en términos del cual viene definido el papel de empleado de banco—, pero cuando entiendo ese papel en esos términos, también entiendo a grandes rasgos las clases de cosas que un hombre tendría que hacer para que se pueda decir de él que es un buen empleado de banco. La diferencia con la idea de un buen padre es que parece como si uno pudiera tener una idea perfectamente clara de la paternidad, y esta no llevara, sin embargo, por sí sola a una comprensión de las clases de cosas que convierten a alguien en un buen padre. La explicación de esta diferencia radica en parte en el hecho de que la idea de paternidad que podemos comprender sin comprender los criterios es una idea de la paternidad meramente como relación biológica; pero no puede ser meramente *esa* idea la que aparece en la

[3] Cfr. G. Cohen, «Beliefs and Roles», *Proceedings of the Aristotelian Society*, vol. 67 (1966-67).

noción de un buen padre. Si sólo tenemos presente la idea de padre como procreador, ni tan siquiera está claro qué es lo que podríamos querer decir al llamar a alguien «un buen padre» —a no ser que queramos decir que era bueno cuando se hizo padre.

Tampoco avanzamos necesariamente hacia nociones evaluativas si nos limitamos a introducir una referencia a instituciones sociales, tales como el matrimonio. «Cuñado» es una bien definida relación de parentesco que envuelve una referencia al matrimonio, pero la idea de ser un buen cuñado es una idea que, en nuestra sociedad, carece de todo contenido. El contenido de «un buen padre» viene determinado por la relación biológica junto con el cumplimiento de ciertas responsabilidades que, en nuestra clase de sociedad, están adscritas a personas que se encuentran en esa relación. Sólo cuando existen tales responsabilidades podemos decir que «padre» no se refiere meramente a la relación biológica, sino a un papel; y el papel puede, por supuesto, ser ejercido en casos excepcionales por alguien que no está en esa relación biológica, como cuando se dice de alguien que ha sido un padre para un niño huérfano. La diferencia con respecto a «empleado de banco» está entonces en lo siguiente: que no existe una noción de «empleado de banco» que no envuelva una referencia a responsabilidades, y en que el término se refiere a un rol del que sólo se puede dar razón en relación con instituciones sociales que al dar el rol imponen ciertas funciones y deberes, mientras que bajo el concepto del rol de la paternidad está el concepto más restringido de paternidad como una pura relación natural que en nuestro tipo de sociedad se entiende como la base de ese tipo de rol.

Si un hombre es un ardiente jugador de criquet, o se ha hecho jugador profesional de criquet, entonces es obvio que le interesará su carrera como jugador de criquet y las evaluaciones de esa carrera por las personas entendidas; si resulta mediocre, ha fracasado. Es igualmente obvio que un hombre que sin más pretensiones tome parte en una partida de criquet un domingo por

la tarde para pasar el rato, tenderá a preocuparse menos; la evidente verdad de que como jugador de criquet es un mediocre significa muy poco para él —podría decir que él no es un jugador de criquet, sino solamente alguien que gusta ocasionalmente de jugar al criquet. Un hombre que fuera un empleado de banco muy indolente, que apenas hiciera lo que se esperaba de él, podría igualmente, si bien de una forma más bien distinta, haberse disociado a sí mismo del cumplir o no cumplir bien en esas actividades. Puede que deteste el banco, desprecie las actividades bancarias y que sólo se preocupe por sus amigos y por sus macetas de crisantemos. Al banco sólo lo ve como un medio de vida; y no se ve a sí mismo, en ningún sentido importante, como empleado de banco. Sin embargo, difícilmente podría decir que *no es* un empleado de banco (en la realidad), o si lo hiciera, podría correr el riesgo de estar refugiándose en la fantasía; pues buena parte de su vida está ligada a ese rol que, le guste o no le guste, tiene. Las actitudes desfavorables y la falta de consideración de que es objeto por sus superiores en el banco, aunque haya logrado que le importen muy poco, es algo que en cualquier caso no puede despachar (como puede hacerlo el hombre que ocasionalmente juega al criquet) alegando que se basan en un malentendido: la relación contractual de su oficio lo pone en una situación en la que no puede esperar otra cosa. Aunque haya alcanzado una cierta estabilidad en esta situación, irónicamente tal vez, la situación no es una situación satisfactoria, y si le tocara la lotería muy bien podría (si esto no exigiera demasiado coraje) desembarazarse de su rol e irse del banco.

En contraste con estos prototipos simples, he aquí otro: el del soldado reclutado contra su voluntad. Su vida, probablemente mucho más penosa que la del empleado de banco, *podría* implicar un estado de ánimo aún más incómodo y ambiguo —si, por ejemplo, no está seguro de si, siendo soldado como es, debería tratar de hacerlo bien como soldado. Pero puede que su estado de ánimo sea también más simple que todo eso si lisa

y llanamente da en pensar que «soldado» es un título que le incumbe solamente porque le ha sido aplicado a la fuerza desde fuera; si nada le importan las evaluaciones que acompañan a este título; si mira la desconsideración de que es objeto por parte de sus superiores como pura fuerza externa igual que la coacción que lo ha puesto en el ejército; y si siente que el único límite de su voluntad de hacer lo menos posible es el miedo al castigo. Lo directo de esta actitud puede que se pague al precio de una cierta desesperación, pues naturalmente (como ocurre en *Catch 22)* tiene que ir acompañada del sentimiento de que, como los hechos brutos cantan, el entorno está loco.

En otros casos, un hombre puede llegar a disociarse a sí mismo de un rol y de las evaluaciones que se le aplican, con los que antes se había identificado a sí mismo de todo corazón o de forma no reflexiva, por no poderse poner a hacer algo que se espera que haga en ese rol. En circunstancias favorables puede que sea capaz de dejar el rol. Si esto no es posible, puede que se sienta obligado a abstenerse de una u otra manera de hacer lo que podría esperarse de alguien en su rol: en un caso político extremo por desobediencia o subversión encubiertos. Se dice de ciertos generales alemanes que durante la guerra estaban horrorizados por las prácticas de Hitler, que durante mucho tiempo no se habían atrevido a ir contra él en consideración del juramento que, como oficiales de la *Wehrmacht*, habían prestado de obedecerle. Uno puede preguntarse, viendo las cosas desde fuera, cómo un juramento prestado a un hombre que estaba entonces en las condiciones en que estaba y que gobernaba por tales medios pudo considerarse como indisolublemente ligado con los deberes de uno, por más que se tratara de los deberes de uno *como oficial alemán* —esto es considerar que los juramentos tienen, no ya santidad, sino poderes mágicos. Pero supongamos que pudiera establecerse que la obediencia al juramento formaba efectivamente parte de los deberes de uno como oficial alemán; entonces, lo que necesitaban los generales era hacerse a la idea de que tenían que hacer cosas contra-

rias a lo que se esperaba que hicieran como oficiales alemanes —es decir, ya no era a título, y en el rol, de *oficial alemán* como habían de actuar. Y eso era una idea a la que muchos de ellos, según parece, encontraron muy difícil hacerse. Para ellos, «oficial alemán» no era justamente algo que ellos eran, sino lo que ellos eran.

Estos ejemplos esquemáticos tienen por objeto ilustrar las formas en que los hombres pueden disociarse a sí mismos de los roles de los que son portadores, roles que llevan consigo ciertas clases de evaluación de sus actividades. Esta disociación puede, en los distintos casos y en diverso grado, ser defendible o indefendible, noble o innoble, prudente o imprudente; pero en cualquiera de los casos es inteligible, y es comprensible cómo un hombre puede formarse una imagen coherente de sí mismo en relación con un rol que tiene, o que tenía o que pudiera pensarse que tiene, imagen en la que no admite que los estándares que acompañan a ese rol sean últimos, básicos o importantes en la evaluación del éxito o excelencia de su vida. Ahora bien, esta posibilidad me parece que representa un esencial contrapeso al hecho que ya hemos señalado, a saber: que estas diversas clases de títulos y roles pueden llevar conceptualmente consigo estándares latos de evaluación de las personas que son sus portadores, al igual que las descripciones de los artefactos pueden llevar consigo estándares de evaluación de esos artefactos. Y aunque los estándares pueden estar, de esta forma, lógicamente fusionados con el título, el título no está lógicamente fusionado con el hombre; con lo que los estándares tampoco están lógicamente fusionados con el hombre. Al hacerse consciente de un determinado título y de su relación con él, un hombre puede negarse a convertir estos estándares en los determinantes de su vida.

De estas posibilidades no se sigue que alguien que, de forma no reflexiva y pasivamente, lleve una vida estructurada por un rol que da por supuesto tenga de hecho que haber *escogido* ese rol, como Sartre (algunos de cuyos temas quedan reflejados en estas observaciones)

parece haber pensado en su periodo existencialista clásico. De ahí la descripción que hace de una condición no reflexiva de esta clase como condición de *mala fe*. Eso exigiría por lo menos que todo hombre tuviera la posibilidad de reflexión y de elección y, además, que esa posibilidad hubiera sido advertida y que se hubiera renunciado a ella. Sin embargo, incluso el primer paso, el de una posibilidad real de reflexión y elección, puede que sea algo que resulte excesivo atribuir a los hombres en algunas situaciones psicológicas y sociales. El punto de vista de Sartre está en conexión con su noción extremadamente no-empírica de libertad, una noción que puede también ayudar a explicar una cierta ambigüedad de su pensamiento entre las ideas de que el hombre no tiene esencia y de que el hombre tiene una esencia que es la libertad.

Si hubiera algún título o rol con el que estuvieran en necesaria conexión determinados estándares y del que, por necesidad, un hombre no pudiera dejar de ser portador ni pudiera disociarse a sí mismo, entonces habría algunos estándares que un hombre tendría que reconocer como determinantes de su vida, por lo menos bajo pena de perder en absoluto la conciencia de lo que realmente es. Existe ciertamente un «título» —por muchas razones es difícil hablar aquí de «rol»— que es necesariamente inalienable, y que es el título de «hombre» mismo. De ahí que sea una cuestión central la de si «hombre» es un concepto que proporcione por sí mismo estándares de evaluación y de excelencia como hombre; pues si lo hace, entonces parece que *tienen que* ser nuestros estándares.

Como se sabe de sobra, han existido filosofías que han sostenido que lo hacía: que, o bien directamente, reflexionando sobre el concepto de «hombre», o bien de forma más indirecta, reflexionando sobre algunas otras verdades supuestamente necesarias relativas a lo que el hombre es, uno podría llegar a la comprensión de lo que un hombre bueno debe ser.

Puede ser útil, por lo menos inicialmente, dividir a estas filosofías en dos clases: en filosofías que hacen y

filosofías que no hacen una apelación a algo trascendente, es decir, una apelación a un marco de referencia que está fuera de la vida humana y del mundo empírico. Podemos empezar con las del segundo tipo y volver después a un ejemplo (un ejemplo religioso) de las del primer tipo.

Estándares morales y la nota distintiva del hombre

Un prototipo del tipo no trascendental puede encontrarse en la filosofía de Aristóteles. Según Aristóteles, hay ciertas características, en particular, ciertas actividades y facultades, que son distintivas del hombre, y la vida del hombre bueno ejemplificará en el más alto grado el desarrollo de estas facultades y actividades. O, dicho con más exactitud, hay un rasgo distintivo del hombre —su capacidad para conformar sus acciones y disposiciones por la razón— que se pondrá de manifiesto en el *más alto grado;* otras de sus potencialidades se realizarán, bajo la capacidad ordenadora de la razón, de una forma equilibrada, y no cada una de forma máxima; se supone que la razón práctica impone coherencia, y reduce el conflicto, entre los deseos de la vida individual (pues el hombre ha de vivir) en sociedad. El objetivo de reducir los conflictos entre deseos, sin suprimirlos desordenadamente, es parte de lo que da un contenido reconocible a la pretensión de que el objetivo de la clase de vida bosquejada en el sistema de Aristóteles es la *felicidad.*

La importancia de la armonización de deseos en Aristóteles y la importancia de la razón práctica en su función de asegurar esa armonización viene ilustrada, si bien en negativo, por su notable fracaso para solventar un problema de reconciliación que, en sus propios términos, tiene que ser importante. La «razón» a la que hasta ahora nos hemos referido es la razón *práctica*, cuyo objeto son las acciones y deseos particulares y que es el fundamento

de lo que Aristóteles (o mejor, sus traductores) denominan «virtudes del carácter», es decir, aquellas disposiciones para la acción correcta que implican motivaciones de placer y de dolor. Está también, sin embargo, la razón *teórica*, la facultad de pensar correctamente sobre cuestiones abstractas de ciencia y de filosofía, que Aristóteles está dispuesto a considerar como una expresión aún más alta de la naturaleza del hombre: la forma más excelente de vida humana es, en consonancia con esto, una vida dedicada en sus justos límites a la pesquisa intelectual. Pues Aristóteles hace ver claramente que puesto que el hombre es un hombre, y no un dios, su vida no puede estar dedicada solamente a esto, y tiene también que tener una vida para la que las virtudes del carácter resultan necesarias.

Lo que Aristóteles no hace, empero —y supuesto su sistema, no puede hacerlo—, es dar una explicación de cómo las actividades intelectuales, la expresión más alta (a su juicio) de la naturaleza del hombre, tienen que ser puestas en relación con las actividades ciudadanas que son reguladas por las virtudes del carácter. La sabiduría práctica no llega a tanto y no podría fijar un «término medio» entre el filosofar o el hacer ciencia por un lado, y el ser un buen ciudadano, un buen padre, etcétera, por el otro. Un rasgo curioso y significativo del sistema de Aristóteles es que se admite que las más altas potencialidades del hombre, al buscar expresión, tienen que entrar en competencia con otras, pero no puede darse una explicación coherente de cómo ha de ser regulada esta competencia.

Esta es una de las debilidades de la tentativa de Aristóteles de educir el bien del hombre a partir de la naturaleza del hombre, apelando para ello a la nota distintiva del hombre, a su inteligencia y capacidad para el pensamiento racional; una debilidad que puede verse en la perspectiva siguiente: que los aspectos puros o creativos de la inteligencia parecería que son la forma más alta de esas capacidades; sin embargo, queda excluida una entrega total a su expresión, y una entrega menos que total no está representada como algo a lo que el pensamiento

práctico pueda llegar racionalmente. Si vemos la dificultad a esta luz, esta debilidad del sistema de Aristóteles puede aparecer como un modelo de un problema más general y bastante real: el problema que podríamos denominar tal vez el «problema Gaugin», con la salvedad de que la etiqueta comporta aquí connotaciones algo especiales y románticas de las pretensiones de *auto*-expresión. Un moralista que desee basar una concepción de la clase de vida correcta para el hombre en consideraciones sobre las facultades superiores y distintivas del hombre, apenas si podrá pasar por alto las pretensiones del genio creativo en las artes y en las ciencias de ser incluido en un lugar preeminente entre tales facultades; sin embargo, le resultará difícil educir de, o incluso reconciliar con, un ideal del desarrollo de la expresión de tal genio, las virtudes y compromisos más cotidianos y domésticos, de los cuales la moralidad ha de dar alguna razón.

Platón, hay que decirlo, vio la mitad de este problema con extrema claridad. La famosa proscripción que se hace en la *República* (un Estado proyectado precisamente como una institucionalización de la moralidad) de todo arte a excepción del narrativo y encomiástico tiene su origen en que tomaba al arte en serio y en que se daba cuenta de forma totalmente correcta de que tanto la vida del artista creador, como los efectos, sobre un público que se lo tomara en serio, de la libre exploración implicada en sus creaciones, era probable que chocaran con las exigencias de estabilidad que implica una sociedad integralmente moralizada. Podemos muy bien rechazar la alternativa que escogió Platón, pero no cabe duda de que tenía razón al abordar el problema en términos de una alternativa: una sociedad protegida contra el cambio moral y social, y una sociedad en la que se permite la libre creación y exploración por medio del arte. La razón por la que (como he dicho) sólo vio la mitad del problema es que permitió, e incluso convirtió en la actividad central de la clase gobernante, la actividad intelectual creadora —y uno puede muy bien preguntarse si esto, a la larga, hubiera resultado ser más compatible

con la clase de estabilidad que deseaba. (Parece que admiraba las realizaciones intelectuales de Atenas aunque deplorando su desorden político y moral, y que admiraba la estabilidad política de Esparta aunque lamentando su militarismo a-intelectual. Sin duda tuvo que habérsele ocurrido preguntarse si era precisamente una casualidad el que esas cualidades estuvieran distribuidas como lo estaban.) Parte de la explicación es que donde nosotros hablamos de «actividad creadora», Platón pensaba ampliamente en términos de descubrimiento: la actividad filosófica estaba a su juicio inherentemente disciplinada por verdades morales *a priori* que estaban aguardando ser descubiertas. Si rechazamos esta imagen de la filosofía moral y admitimos además que en la ciencia natural hay un ineliminable elemento creador, estas actividades naturales pierden bastante de ese aspecto de habitantes nativos de la clase de entorno que Platón les tenía destinado.

No estoy diciendo (pues sólo me parece un generoso sinsentido) que las propias actividades de la investigación científica exijan o presupongan que las personas implicadas en tal investigación tengan que tener valores liberales o humanos, opuestos a un entorno social autoritario. La reciente experiencia sugiere que lo más que tales investigaciones presuponen es un cierto liberalismo en el entorno de la ciencia misma, que puede muy bien coexistir con una indiferencia cínica hacia muchas consideraciones humanas, relativas (por ejemplo) al uso que se hace de los descubrimientos científicos o a las fuentes de financiación de la investigación. Pero esto en sí mismo no es más que otra ilustración de las varias formas en que las exigencias de la moralidad pueden entrar en conflicto con el desarrollo, abandonado a sí mismo, de las aspiraciones creadoras e intelectuales del hombre: nos ha tocado vivir una hora lo suficientemente avanzada en la jornada de la historia como para que podamos ponernos a decir que, dado que la ciencia natural claramente constituye una de las más altas realizaciones de los hombres, su desarrollo puede elevar una indesafiable pretensión a nuestra aprobación moral. Sería tan opti-

mista de nuestra parte el pensar esto como lo era por parte de Platón el pensar que el desarrollo de esas actividades podría coexistir con el grado de restricción social y de no liberalismo que *él* identificaba con la sociedad moralizada.

Esto ilustra una objeción, bastante central, que puede hacerse a la tentativa aristotélica, con referencia a la particular (y plausible) elección del intelecto como nota distintiva. Existen otras objeciones más generales al procedimiento con que trata de educir fines e ideales morales incuestionables a partir de las notas distintivas de la naturaleza humana. Podemos mencionar tres. En primer lugar, se ha introducido ya un palpable grado de evaluación en la selección de la nota distintiva a la que se le da este papel, como son la racionalidad o la creatividad. Si uno abordara sin prejuicios el problema de encontrar características que diferencien al hombre de otros animales, uno podría también, basándose en estos principios, acabar en una moralidad que exhortara a los hombres a hacer fuego; o a desarrollar características físicas peculiarmente humanas; o a tener relaciones sexuales sin mirar la ocasión; o a despojar el entorno y trastornar el equilibrio de la naturaleza; o a matar cosas por diversión.

En segundo lugar, y esto es básico, este planteamiento corrobora la *ambigüedad* moral de las características distintivas del hombre (si bien Aristóteles, aunque su éxito aquí tampoco fuera completo, prestó alguna atención a este punto). Pues si es una nota del hombre el emplear su inteligencia e instrumentos en modificar su entorno, también habrá de ser igualmente una nota del hombre el emplear su inteligencia en ir a lo suyo, y sus instrumentos en destruir a los otros. Si es una nota del hombre el tener conciencia, refleja y plena, de sí mismo como de uno entre otros, consciente de que los otros tienen sentimientos como él, esto es una precondición no solamente de benevolencia, sino (como subrayó Nietszche) también de crueldad: el sádico no se parece más a otros animales que el hombre de afectos naturales, sino menos. Si como imperativo moral supremo propo-

nemos «sé un hombre», es terrible pensar las muchas formas en que puede ser tomado de forma literal.

Aquí parece que damos con una genuina dimensión de la libertad, la libertad de hacer uso o de descuidar las dotes naturales, o de hacer uso de ellas de una manera o de otra: una libertad que por necesidad tiene que cortar la cuerda central de proyectos del tipo del de Aristóteles. Tampoco puede utilizarse la libertad como nota distintiva del hombre, y rehacer de nuevo el proyecto sobre esta base. Pues esta libertad, sin duda, no puede, por su propia naturaleza, determinar una forma de vida frente a otra —como quizás pensó Sartre en virtud de una ambigüedad a la que ya nos hemos referido. Ya que uno podría decir: si existe una forma distintiva de vida, la de «realizar la libertad», entonces tiene que existir también la libertad de rechazarla.

En tercer lugar, y volviendo al caso particular de lo *racional* como nota distintiva del hombre, se registra en este planteamiento una tendencia a adquirir una inclinación maniquea y a poner el énfasis en las virtudes del autocontrol racional a expensas de todo lo demás. No hay razón ninguna por la que *inevitablemente* hubiera de seguirse tal visión de las cosas; aparte de todo lo demás, implica un concepto falso e inhumano de las pasiones mismas como fuerzas causales ciegas o como características puramente animales. Sentirse irremediablemente enamorado es una condición tan distintivamente humana como lo es la aprobación racional de las disposiciones morales de alguien. Pero es fácil ver por qué, en esta dirección, el maniqueísmo acaba resultando tentador. Si la racionalidad y el pensamiento consistente son las notas distintivas preferidas, entonces, aun cuando se admita que el hombre, como un todo, tiene también pasiones, la supremacía del pensamiento racional sobre ellas puede muy bien llegar a ser considerada como un ideal incuestionable. Y esto tanto más cuanto que resulta obvio que la obtención de un cierto control de esa especie es una condición básica para el desarrollo, e incluso, en el caso extremo, para la salud mental. Pero pasar de eso a hacer de tal control *el* ideal es deses-

timar *a priori* casi todas las formas de espontaneidad. Y esto parece que es absurdo.

Todas estas consideraciones vienen a sugerir que la tentativa de fundar la moralidad en una concepción del *hombre bueno* educida de consideraciones sobre las notas distintivas de la naturaleza humana parece condenada al fracaso. Estoy muy lejos de pensar que las consideraciones sobre la naturaleza humana, lo que los hombres son, lo que para los hombres significa vivir en sociedad, no contribuyan a una correcta concepción de la moralidad. Por supuesto que lo hacen: sin esas consideraciones uno carecería en absoluto de toda concepción de la moralidad. En particular, contribuyen a delimitar el posible contenido de lo que podría ser considerado como una moralidad. Y tan obvio como lo anterior es que diferentes concepciones de la naturaleza humana (como, por ejemplo, alguna concepción psicoanalítica) tienen que tener diferentes efectos sobre las ideas que uno se hace sobre determinados requerimientos y normas morales. Y no solamente las ideas científicas o semicientíficas habrán de tener ese efecto, sino también los puntos de vista que se sostengan en la filosofía de la mente. Así, una comprensión adecuada de la naturaleza de las emociones habría de tener un efecto desalentador sobre los puntos de vista maniqueos en relación con su control, y las consideraciones filosóficas sobre la naturaleza, y por supuesto que sobre la existencia, de algo llamado la *voluntad*, deben tener un efecto directo sobre las moralidades que encuentran en el ejercicio de la voluntad (contra los deseos, por ejemplo) una clave central para el valor moral.

Aunque todo esto es verdad, y aunque existen límites muy definidos que no podría sobrepasar lo que pudiera ser considerado, de forma comprensible, como un sistema de moralidad humana, no hay ruta directa que vaya de las consideraciones sobre la naturaleza humana a una moralidad única o a un único ideal moral. Todo sería más simple si hubiera menos cosas, y menos cosas distintivamente humanas, que los hombres pudieran ser; o si los caracteres, disposiciones, instituciones sociales

y estados de cosas que pueden, de forma comprensible, tener valor para los hombres, fueran todos, en desarrollo pleno, consistentes los unos con los otros. Pero no lo son y existen buenas razones para que no lo sean: buenas razones que, también ellas, emergen de consideraciones sobre la naturaleza humana.

Dios, moralidad y prudencia

Hemos hecho una distinción hace un momento dentro de las concepciones que tratan de educir la noción de un hombre bueno de consideraciones sobre la naturaleza del hombre: las que ponen al hombre en un marco de referencia trascendental y las que no lo hacen. Después de haber dicho algo sobre la segunda clase, vamos a pasar a examinar un ejemplo de la primera. Y a la par que hacemos esto, no estará de más discutir una cuestión que es en sí misma importante con independencia del tema que nos ocupa —las relaciones entre la moral y la prudencia.

Un rasgo capital de esta clase de teorías es que tratan de señalar, en términos del marco de referencia trascendental, algo *a lo que* el hombre está destinado: si el hombre entiende adecuadamente su papel en el plan básico de las cosas, se dará cuenta de que existen determinadas clases de fines que son propiamente los suyos y que él debe realizar. Una forma arquetípica de tal concepción es la creencia de que el hombre fue creado por un Dios que también espera ciertas cosas de él.

Una dificultad central con que se enfrenta esto es la cuestión de qué propiedades de Dios se supone que justifican la pretensión de que nosotros debemos satisfacer sus expectativas. Si se trata de su poder, o del mero hecho de que nos ha creado, las analogías con los reyes y padres humanos (frecuentemente empleadas en relación con este tema) nos dejan con la impresión de que hay muchos reyes y padres que no deben ser obedecidos. Y si se nos

insiste en que Dios tiene un poder infinito y en que ha creado todo, replicamos que no es evidente que su omnipotencia y su carácter de creador de todas las cosas sean más dignos de ser obedecidos, sino que, meramente, resultan más difíciles de ser desobedecidos.

Si entonces se dice que, además de estas otras propiedades, Dios es bueno, la objeción siguiente (como en el caso de Kant) es que esto ya implica un reconocimiento de lo que es admirable y valioso, un reconocimiento de la clase que la apelación a Dios se suponía que tenía la función de asegurar.

Tales argumentos, que son muy comunes, pueden ser entendidos como un ataque a la idea de que uno pudiera resolver de forma puramente deductiva y *a priori* la cuestión de la vida exigida al hombre, a partir de una descripción suya *como creado por Dios*. Entendidos así, me parece que son concluyentes. Pero tales argumentos tienden a implicar una ambición más alta —la de probar que, aun cuando se demostrara la existencia de Dios, ese hecho no proporcionaría en principio ningún *motivo* aceptable o apropiado para la conducta moral, de una clase con la que no se podría contar de otra forma. En este papel también, los argumentos han sido objeto de una muy amplia aceptación, de modo que prácticamente parece una perogrullada en filosofía la idea de que si Dios existiera, esto, para un pensador moral y con las ideas claras,en nada cambiaría las cosas en relación con la situación de la moralidad. Los orígenes de esta concepción se remontan a una famosa discusión en el *Eutifrón*, de Platón, pero en su forma moderna se debe casi por completo a Kant. Debemos en particular a Kant una clara exposición de los supuestos en los que se basa —supuestos relativos a la *pureza* esencial de la motivación moral. Estos supuestos tienden a penetrar y han penetrado en buena parte del pensamiento moral, de modo que su influencia e importancia van mucho más allá del presente tema de la moralidad religiosa. La equivocación en que incurren es también importante.

El argumento, en su forma más simple, es más o menos como sigue. O los motivos de uno para seguir la palabra

moral de Dios son motivos morales, o no lo son. Si lo son, entonces uno está ya provisto de motivaciones morales, y la introducción de Dios no añade nada extra. Pero si no son motivos morales, entonces serán motivos de clase tal que no pueden en absoluto motivar la *moralidad* de forma adecuada: en particular es probable que sean motivos prudenciales, una posibilidad cruelmente pintada por muchos evangelistas (tanto de la fe como de la incredulidad) en términos de fuego del infierno. Pero nada motivado por consideraciones prudenciales puede ser una acción genuinamente moral; la acción genuinamente moral tiene que estar motivada por la consideración de que es moralmente correcta, y no por otra clase de consideraciones en absoluto. Así, teniendo en cuenta todo ello a la vez, llegamos a la conclusión de que, en este tema, la apelación a Dios, o no añade nada en absoluto, o añade algo que es un error.

Esta clase de argumentación suscita dos cuestiones en relación con la moralidad y la motivación. La primera es la de si no existen realmente otros tipos relevantes de motivación que la moral o la prudencial, es decir, la cuestión de si la distinción entre lo moral y lo prudencial es exhaustiva. La segunda es la de si un curso de acción o una concepción no podrían ser morales siendo de algún modo, al mismo tiempo, prudenciales, es decir, la cuestión de si la distinción es excluyente. Vamos a ocuparnos en primer lugar de la segunda cuestión. ¿Es esencial para la moralidad distinguir totalmente lo moral de lo prudencial?

Aquí necesitamos hacer algunas distinciones. Es ciertamente verdad que es esencial para la moralidad el que *en cierto nivel* se trace una distinción entre lo moral y lo prudencial. En el plano más primitivo, es claro que toda moralidad tiene que aplicar esta distinción, o algo similar, a las acciones y proyectos; tiene que ser capaz de distinguir entre acciones y proyectos que son *egoístas* y que están al servicio de la gratificación o la seguridad del agente a expensas de los otros, y los que tienen en cuenta los intereses de los otros. Si no se hace alguna distinción de este tipo, no hay consideraciones morales

en absoluto. Es claro que la moralidad religiosa que estamos discutiendo, aunque expresándonos con crudeza, puede observar esta distinción en un nivel primitivo: con respecto al mundo secular aprobará proyectos y acciones que tienen en cuenta el interés de los otros, y desaprobará los comportamientos egoístas.

Sin embargo, puede pensarse que el nivel en el que se está trazando la distinción es demasiado primitivo y que es preciso hacer extensiva la distinción a la discriminación entre motivos, y no quedarse sólo en la mera clasificación de comportamientos y acciones intencionales. Así, uno que da dinero con destino a obras benéficas meramente para mejorar su reputación ante el Club Rotario, para aligerar sus propias culpas, no actúa más moralmente que si hubiera empleado el dinero en sus propios gustos. (El ejemplo ilustra por qué era oportuno hablar de «acción intencional» y distinguir ésta de la cuestión del motivo. El empresario autointeresado que extiende un talón destinado a los fondos para la lucha contra el hambre, lo hace intencionalmente, y su intención es que el dinero se destine a aliviar el hambre: si con su acción se alivia el hambre, esto no será un *accidente* en relación con sus pensamientos al actuar así. La cuestión es que su motivo no era la preocupación por aliviar el hambre, sino la preocupación por su propia reputación o comodidad.)

Si decimos, como muchos dirían, que el hombre que actúa así, no actúa más moralmente que uno que emplea su dinero en sí mismo, no se seguirá que lo que hace no es *mejor* que lo que hace el hombre simplemente egoísta; pues el hambre, cabe esperarlo, quedará aliviada, y esto es mejor que el que se expenda otra coctelera o un aparato de televisión. Ni tampoco puede ciertamente seguirse que nuestra aprobación se limite meramente al acto y que en *ningún* sentido se haga extensiva al agente; pues podemos decir de él que intencionadamente ha hecho algo que está bien hecho, y esto, sin duda alguna, constituye algún tipo de aprobación, aunque sea limitada, del agente. Pero posiblemente la cuestión sea que no lo aprobamos *moralmente*. La idea tiene, desde luego,

gran interés, pero no conviene llevarla demasiado lejos. Pues si insistimos en que actuar moralmente es actuar por una motivación moral, podemos muy bien sentirnos tentados a añadir a eso la proposición aparentemente inocua de que todo lo que puede importar desde un punto de vista moral es que la gente actúe moralmente, y sacar entonces la conclusión (correcta a partir de esas premisas) de que desde el punto de vista moral dos situaciones, cualesquiera que éstas sean, de motivación autointeresada son indistinguibles, y tiene que ser imposible desde un punto de vista moral preferir la una a la otra. Esto es absurdo. Tal vez no sea lógicamente absurdo, pero es moralmente absurdo —el absurdo moral del puritano de que la única propiedad moralmente relevante del mundo es la cantidad de honradez que contiene. Pero como esto es moralmente absurdo (o mejor, como es obvio que no es incoherente considerarlo moralmente absurdo), se sigue que tiene que haber otra cosa que *sea* la lógicamente absurda, y que es la concepción de la moral de la que se sigue que esta es la única posición moral que es sostenible de forma coherente.

¿En qué está, pues, el meollo y contenido de decir que no aprobamos *moralmente* al que autointeresadamente hace una donación para una obra benéfica, o que, aunque lo que hace está bien, no actúa moralmente? ¿Con qué motivaciones, en primer lugar, estamos contrastando las motivaciones de este hombre? Algunos, como Kant y R. M. Hare, han puesto el énfasis en el contraste que existe entre la actuación de este hombre y la actuación *por principios;* dicho sin más precisiones, entre la actuación de este hombre y el hacer algo precisamente porque uno debe hacerlo. Otros, como Hume, han subrayado el contraste que existe entre esa actuación y el hacer algo porque uno se preocupa desinteresadamente por la situación que se supone que la acción de uno va a cambiar o porque se preocupa por las personas implicadas. Dejemos de lado las notables diferencias entre estas dos formulaciones, para considerar lo que tienen en común: que el hombre que tiene una motivación moral para hacer cosas de tipo no autointeresado, tiene

una disposición o motivo general para hacer cosas de este tipo; mientras que el hombre autointeresado no tiene ese motivo estable, pues nunca será más que una afortunada coincidencia que lo que beneficia a otros sea también lo que, compuestos los criterios del simple autointerés, le beneficia a él. Y esto, sin duda alguna, y como decía Hume, ha de tener algo que ver con la *clave* de la selección que hacemos de ciertos motivos para la aprobación *moral:* nos es de importancia tener personas que tengan una tendencia general a estar dispuestas a poner los intereses de otras personas por encima de los suyos propios.

Tal vez merezca la pena hacer notar, de pasada, que una de las (numerosas) ventajas del énfasis que pone Hume en este tema, con su acento en la simpatía y la sensibilidad ante la situación de los otros, en contraste con el énfasis que pone Kant en la actuación por principios, es que introduce una semejanza entre las clases de razones que uno tiene para hacer cosas por otros, y las clases de razones que uno tiene para hacerlas por uno mismo. A pesar del carácter mecánico del sistema psicológico de Hume, da algún significado comprensible a la idea de que preocuparse por el dolor de otros es una extensión de la preocupación por el de uno mismo: lo segundo es desde luego, una condición necesaria de lo primero, y no se hace ningún problema (no tendría por qué hacerse) de por qué un hombre que se preocupa por los otros no puede estar también razonablemente preocupado por sí mismo. Sin embargo, bajo el énfasis kantiano, esto aparece de pronto como un problema, pues el actuar con la mira puesta en los propios intereses de uno es actuar por una clase de motivos que nada en absoluto tienen que ver con la moralidad y que, obviamente, son ajenos a ella. Como es de presumir que estamos obligados a maximizar la acción moral, parece que habrían de seguirse extremos de autonegación, como algo derivado, por supuesto, del concepto mismo de moralidad. En el mejor de los casos, hacer lo que, sin más, uno desea hacer, constituirá una desviación desordenada y probablemente culpable con respecto al punto

de vista moral. Para enfrentarse a este problema, la tradición kantiana presenta un conjunto de «deberes para consigo mismo», el reconocimiento de los cuales deja libre a uno para hacer por razones morales algunas de las cosas que haría de todos modos. Este absurdo artificio es justamente el resultado de tratar de ajustar a una visión algo más razonable de la vida humana las embarazosas consecuencias que se siguen de sostener tres cosas —que la moralidad se refiere por encima de todo a la motivación, que la motivación moral es motivación de principio, y que el punto de vista moral ha de ser ubicuo. Para evitar estas embarazosas consecuencias bastaría con abandonar una cualquiera de estas proposiciones; pero existen buenas razones para abandonar las tres.

Volvamos, después de esta disgresión, a nuestro moralista religioso. Vimos que podria fácilmente trazar una distinción, aunque primitiva, entre lo moral y lo prudencial. Pienso que ahora podemos ver que puede incluso establecer una distinción más refinada en el plano de la motivación. Hemos sugerido que uno de los aspectos cruciales (aunque tal vez no el único) de la distinción entre motivación moral y motivación autointeresada era que nos permitía seleccionar disposiciones generales a hacer cosas del tipo no-autointeresado. Pero las disposiciones de este tipo serán algo que ciertamente admitirá incluso el hombre de moralidad religiosa más cruda. Puede quizás que lo que su Dios desea es que los hombres sean sensibles a los sufrimientos de todos y cada uno de los otros y que cada uno actúe en interés de los otros porque así lo siente. Así, muchas veces, las personas con esta creencia, si hicieran lo que Dios quiere que hagan, actuarían por motivos humanos ordinarios que la mayoría de la gente (que no sea kantiana) consideraría como motivos en sí mismos morales. Si estos fallan, o la tentación de actuar de forma egoísta es fuerte, entonces tal vez los pensamientos de este rudo creyente se vuelvan por un momento al fuego del infierno, y esto fortifique la disposición a hacer cosas de tipo no-autointeresado (en el sentido mundano de «autointeresado»,

claro está). Y si bien su acción es en este sentido una acción prudencial, no es prudencial en el sentido en que es esencial para el concepto de la moralidad que lo prudencial se distinga de lo moral. Y por supuesto, hay una especial razón por la que sus acciones, aunque prudenciales, no son egoístas, a saber: es de presumir que no exista forma efectiva de tener como meta la salvación *a expensas de los otros.*

De hecho, es bastante carente de realismo el querer imponer a nuestro moralista religioso (o a algún otro) una distinción excluyente entre lo prudencial y lo moral. Dejando de lado los efectos más generales de la simpatía a los que se refirió Hume: ¿qué decir de alguien que hace algo en interés de otro, y en desventaja propia, porque ama a esa persona; o, incluso, está enamorado de ella; o la admira; o la respeta; o porque (después de todo) es un miembro de la familia? Ninguna de estas razones para obrar tienen que ser razones morales en ningún sentido exigente o purificado del término; tampoco son razones prudenciales. Ni tampoco pertenecen a la tercera clase de motivos que la filosofía moral ha permitido alguna vez, los de la *inclinación*, es decir, el hacer una cosa porque a uno le gusta hacerla. Es claro que la lista de ejemplos podría extenderse indefinidamente para incluir un gran número de formas especiales de relación en las que una persona puede encontrarse con otra. Es un grotesco producto de la teoría y de un moralismo obstinado el suponer que lo «moral» y lo «prudencial» se reparten de forma adecuada el campo de los motivos o razones justificables que un hombre puede tener para hacer algo: por el contrario, dejan fuera casi todo. Necesitamos desde luego algo más allá de y por encima de estas motivaciones particulares o más específicas, precisamente porque son particulares y específicas, y especialmente porque la relación particular que puedo tener con otra persona puede ser de carácter hostil, y es del todo probable que no se divise en el horizonte otra relación más, benéfica, también de carácter particular, que sea capaz de inhibir la conducta destructiva. Así, es obvio que necesitamos, allende y por encima de las

motivaciones particulares, motivaciones generales para controlar y regular esas motivaciones particulares; y la cosa más general allende y por encima es la moralidad. Afortunadamente para la humanidad, no tenemos que esperar que sean estas consideraciones generales las que motiven todo lo deseable. Algunas de nuestras acciones honestas no provienen de ese motivo que los cristianos malinterpretan como nuestro amor al prójimo, a todo el mundo, sino justamente de que amamos a alguien.

Ahora bien, el moralista religioso puede considerar que esas exigencias generales tienen su fuente en una relación particular, en la relación con Dios, y esta relación puede ser representada como una relación de amor, de reverencia, de respeto o cualquiera sea el término que se considere adecuado en esta ardua tarea semántica. Y tendrá razón al resistirse a que esta relación sea categorizada o bien como moral o bien como prudencial en su trascendencia práctica. Pienso, no obstante, que ha de guardarse cuidadosamente de decir que ésta es una actitud hacia Dios que *debe* tener todo el que conoce lo que es Dios, pues esto podría, en el sentido en que Kant insistía en ello, volver a dar a la moralidad prioridad sobre Dios. Tendrá que decir más bien que esta es una actitud que todo el que conoce lo que es Dios tendrá inevitablemente. Dios es alguien cuya palabra exige una aceptación incondicional. Por supuesto que decir esto no es decir suficiente todavía; pues compatible con esto es que Dios sea un hipnotizador indefectible. Así, el creyente habrá de proceder como siempre por negación y analogía, y decir que no es eso, sino que se parece *más* a un padre bondadoso, etc. Por mi parte dudo de que en último término vaya a ser capaz de dar una explicación coherente en absoluto. Pero esto se debe a las dificultades de la fe en Dios, y no a algo que radique en la naturaleza de la moralidad. No creo que sea muy correcto decir —y estos argumentos han tratado de demostrarlo—, como muchos hacen, que aunque Dios existiera, eso no podría constituir ninguna razón especial y aceptable para suscribir la moralidad. Si Dios existiera, muy

bien podría haber razones especiales, y aceptables, para suscribir la moralidad. El problema es que la tentativa de formular esas razones en forma algo mejor que la de un paupérrimo bosquejo, choca con la imposibilidad de pensar coherentemente sobre Dios. El problema de la moralidad religiosa proviene no de que la moralidad sea ineludiblemente pura, sino de que la religión es irremediablemente ininteligible.

¿Sobre qué versa la moralidad?

Estas últimas discusiones nos han introducido de soslayo en ciertas cuestiones tales como a qué se refiere la moralidad y cómo hay que delimitar lo «moral». Mucho se ha escrito sobre esto, en busca de algunos criterios para distinguir lo moral de lo no moral. Estas cuestiones son examinadas por G. J. Warnock en un libro admirablemente conciso, lúcido y lleno de fuerza, *Contemporary Moral Philosophy*[4]: insiste Warnock con razón en que buena parte de esta discusión ha tenido lugar en el vacío, en busca de un criterio que pudiera proporcionarnos *una* forma de distinguir entre lo moral y lo no moral, tal vez una forma que guardaría una oscura similitud con la forma en que nosotros, ahora, hacemos esa distinción, pero que no arroja ninguna luz sobre, o no se ve guiada por la cuestión, evidentemente más básica, de la *finalidad* de esa distinción, la cuestión de qué interés tiene trazar una línea que divida acciones, conductas, motivos y razones humanas. Algunas de las observaciones que hicimos en la sección anterior se refieren a este punto.

Daré por supuesta —de hecho ya la he dado por supuesta antes— una conclusión a la que llega Mr. Warnock en su discusión y que tiene sin duda que ser correcta, la conclusión de que toda delimitación, con sentido, de lo moral tiene que hacer referencia al *contenido* de los juicios, comportamientos, principios, y todo lo demás

[4] Cfr. ahora su *The Object of Morality*, Londres, 1971.

que se describe como «moral». Tal vez pudiera sorprender a las personas legas en filosofía moral que alguien haya podido suponer alguna vez que esto no fuera así; pero de hecho se ha sostenido, y con frecuencia, que las ideas morales (en contraste con las no morales) pueden ser identificadas sin referencia a su contenido, por algunas consideraciones tales como que son máximas prácticas que son enteramente universales o que son máximas prácticas que son reconocidas como máximas que prevalecen sobre otras máximas prácticas. El motivo de estas maniobras tan poco plausibles ha sido, como siempre, la insistencia en la distinción hecho/valor. Pues supongamos que alguien introduce alguna referencia al contenido de lo moral y dice (por ejemplo) que las concepciones morales se refieren esencialmente al bienestar humano, donde el término «bienestar» tiene algún contenido propio y no significa meramente «lo que uno piensa que debe acontecer a los seres humanos»: entonces, el rango de las concepciones morales posibles queda severamente limitado tanto por los hechos como por la lógica, contrariamente a la distinción hecho/valor. Este motivo para iniciar tales maniobras difícilmente las hace más atractivas que las consecuencias que tienen.

La cuestión que quiero discutir aquí no es, pues, este asunto general, sino los méritos del criterio parcial que acabo de mencionar, el criterio de una referencia al bienestar humano como nota de una posición moral; hecha por Mrs. P. R. Foot y otros, esta propuesta es bien vista por Mr. Warnock. Para que pueda ser defendible, es esencial que ese test se aplique en un plano muy general y que «bienestar» se interprete en términos muy generales —aunque no, por supuesto, en términos tan generales que se convierta en vacío.

El primer punto queda claro si se tiene en cuenta que ciertamente uno no podría aplicar tal test a la motivación de las *acciones particulares* y conseguir siempre resultados correctos: debe tratarse más bien de que si uno aprueba moralmente cosas más generales como proyectos, instituciones, disposiciones, clases de motivos, etcétera, entonces uno tiene que suponer que esos pro-

yectos, instituciones, etc., están en cierta manera al servicio de la realización de alguna clase de bienestar humano. El segundo punto es que, ciertamente, «bienestar» no puede ser entendido, con este fin, con la mera significación de que las personas implicadas en estas situaciones consiguen lo que de hecho desean o de que (lo que no tiene necesariamente que ser la misma cosa) les gusta el resultado. Pues, sin duda, tiene que ser posible admitir como puntos de vista morales (aunque los utilitaristas se inclinarían a pensar que se trata de puntos de vista morales equivocados) concepciones que sostienen que con frecuencia la gente desea y gusta de cosas desacertadas.

Tampoco está claro, aunque esto puede ser más cuestionable, que «bienestar» en este sentido pueda ser identificado con *felicidad*. Ciertamente, si pensamos que la satisfacción o el contento son condiciones suficientes, pero no necesarias, de la felicidad, esto no valdrá para el presente propósito; pues tenemos que ser capaces de considerar como puntos de vista morales (aunque los cínicos los considerarían erróneos) concepciones que deploran la satisfacción y el contento, por lo menos si se obtienen a un nivel muy bajo de conciencia y de actividad. Incluso si eliminamos esta referencia y nos negamos a considerar «felices» a los que están contentos en algún estado bovino, narcotizado o de sumisión —como Aristóteles y, con una noble inconsistencia, J. S. Mill se negaron a considerarlos— e insistimos en un nivel más activo y reflexivo de felicidad, todavía no está claro que tengamos lo que la tesis requiere. La idea de la felicidad de un hombre, y por supuesto la noción menos episódica de que un hombre está siendo *un hombre feliz*, tiene sin duda *algo* que ver con que ese hombre no sufre; o con que no sufre demasiado; o en el límite, con que no sufre de la forma más dolorosa, pues muy bien podríamos decir de un hombre en una condición física penosa que, sin embargo, es feliz si se sobrepone a la desesperación, a la autoconmiseración, etcétera, y conserva interés y placer en otras cosas —y está, en algunas ocasiones por lo menos, moderadamente *alegre*. De la idea de algunos

filósofos de la Antigüedad de que la virtud era suficiente para la felicidad y de que el hombre bueno podría ser feliz en el potro de tortura se pensó, y no sin razón por supuesto, antes, después y sin duda también en su época, que implicaba una paradoja. Pero si la felicidad es, en último extremo, incompatible con un sufrimiento excesivo o con un sufrimiento demasiado global, puede haber tal vez concepciones, que haya que admitir como morales, que rechacen la idea de que la felicidad sea el hilo conductor de nuestros proyectos y designios. Se cuenta de Lutero que, cuando alguien propuso la *Glücklichkeit* (felicidad) como fin de la vida humana, rechazó violentamente la idea diciendo «leiden, leiden, Kreuz, Kreuz» («sufrimiento... cruz...»). Podría tratarse de la idea de que el pecado del hombre y su distancia con respecto a Dios es tal, que sólo una vida de penitencia y reconocimiento de la propia maldad y de la maldad de todos podría responder de forma adecuada a la situación. Tal concepción, sin duda alguna, deploraría las instituciones, proyectos, formas de vida, que tendieran a eliminar esos sufrimientos, los más básicos del hombre, y puede que aprobara de forma contingente o periférica solamente los que aliviaran sufrimientos menos espirituales.

A esto podría replicarse que eso sólo demuestra que Lutero ponía el bienestar del hombre en otra parte, en la eventual reconciliación con Dios, y que los sufrimientos aquí eran solamente medios para la felicidad en otro sitio. De este modo, la felicidad sigue siendo lo importante, aunque se trata de la felicidad en otra parte. Pero esta formulación, por lo menos en la limitada medida en que entiendo la concepción de Lutero, me parece que la desvirtúa de una manera esencial, y también quizás a algunas otras concepciones protestantes. La cuestión es que no hay *medios* a disposición del hombre para una reconciliación con Dios y que no puede concebirse que exista un conjunto de proyectos humanos capaz de asegurar este resultado —el abismo es demasiado grande, y solamente hay un signo de esperanza, Jesucristo, de que la gracia de Dios socorrerá nuestra indignidad. El creyente

obedecerá la voluntad de Dios lo mejor que pueda en su miserable condición, y ha de retener la conciencia de esa condición; pero no *para* asegurarse tanto para sí como para otros esa salvación, lo que en el mejor de los casos sería una insensata esperanza; y si es rechazado, no tiene derecho a quejarse.

A mi juicio, esta clase de concepción protestante podría ser considerada como una concepción moral —pues ciertamente su propósito es explicar la situación del hombre en relación con la clase de conducta que se espera de él, y también habla de lo que, según su modo de ver las cosas, es lo central en relación con la felicidad del hombre. Pero a la felicidad se la ve tan lejos, y al hombre tan alienado de la fuente de la felicidad, que sería una distorsión pensar que la meta que se propone esa moralidad es la felicidad del hombre: la meta es más bien que la vida *refleje*, en el sufrimiento y en la obediencia, la deplorable condición humana. Y esto es sin duda una concepción muy distinta de la que se ocupa, de forma directa, de la felicidad, y sin embargo, sería avieso negarle por ello la denominación de moralidad. Uno podría tal vez decir que las moralidades con una dimensión trascendental tienen una mayor libertad lógica, en relación con su contenido, que las moralidades que no tienen tal dimensión. En esa pintura trascendental hablan del papel y la condición general del hombre en unos términos que pueden hacer parcialmente inteligibles como posición moral actitudes que resultarían bastante opacas si se nos ofrecieran en un marco de referencia puramente secular.

Pero incluso en un marco de referencia secular puede que resulte posible encontrar posiciones morales para las que, igualmente, la «felicidad» constituya sólo una caracterización muy pobre de sus preocupaciones centrales. Así, ciertas posiciones románticas que se expresan en términos de una libre respuesta a la vida, o de «sinceridad» en relación con los propios impulsos, incluyendo los destructivos, o en términos de la significación de experiencias límites, es posible que sólo queden muy pobremente descritas diciendo que tienen una idea muy

particular de la *felicidad* del hombre. Puede ser que algunas de estas posiciones hablen menos del marco general de referencia de la moralidad que de ciertos *ideales* personales, que por supuesto entran en la moralidad en el sentido de que, para aquellos que son sensibles a tales ideales, constituyen un modelo de vida para ser vivido y al que se atribuye una especial importancia, pero se preocupan menos de la cuestión de qué clase de reglas, instituciones, disposiciones, etcétera, se requieren en la sociedad como un todo. Y esto suscita temas de gran alcance, ya que la relación entre los ideales personales y las normas sociales generales es también un importante tema moral.

Sería una necedad tratar de discutir estos últimos temas en términos muy generales: las posiciones en cuestión necesitan ser expuestas y discutidas con detalle y eso es algo que no podemos hacer aquí. Pero parece una cuestión abierta la de si alguna de esas concepciones no cortan, propiamente hablando, la conexión con la felicidad como foco de la actividad moral humana. Una pregunta central que hay que hacerse al considerar esto habrá de ser siempre, a mi juicio, la de en qué medida la posición moral de que se trate hace alguna apelación trascendental de alguna clase, aunque sólo sea de forma tácita o sólo en vestigios. E incluso donde no hay una apelación en el sentido de una referencia, tal como hace el moralista religioso, a algo fuera de la vida humana que proporciones de algún modo un modelo para esa vida, sin embargo, puede haber una apelación a algo que esté *ahí* en la vida humana, y que haya de ser descubierto, a lo que uno haya de entregarse, que haya de ser seguido, incluso ignorando gravemente los resultados.

Como indica la última frase no estoy hablando aquí de algo como la moralidad marxista, que no cae dentro de lo que ahora nos ocupa, ya que lo que preocupa a esa moral es de forma casi directa la cuestión de la felicidad última. El bienestar del hombre, que se aguarda tras la destrucción del capitalismo y la eliminación de la explotación, tal vez no consista solamente en que los hombres vayan a ser más felices, pero ciertamente que

incluye también eso. Más bien, lo que tengo *in mente* es, por ejemplo, algo indicado por una frase de D. H. Lawrence en su espléndido comentario a la complaciente exhortación moral de Benjamin Franklin «descubrid vuestro impulso más profundo y seguidlo». La noción de que *hay* algo que es el impulso más profundo de uno; de que, más que de tomar una decisión, de lo que se trata aquí es de hacer un descubrimiento; y la noción de que uno se confíe a lo así descubierto, aunque no sepa a dónde va a conducirlo —estas son aquí las cosas importantes. La combinación —descubrimiento, confianza y riesgo— es central en esta clase de posición, como también lo es, por cierto, en el estado de enamoramiento. Uno siente incluso la tentación de hallar, entre los muchos legados históricos del protestantismo al romanticismo, un paralelismo entre esta combinación y el par tan importante para Lutero: obediencia y esperanza. Las dos posiciones establecen una conexión esencial entre sumisión e incertidumbre; las dos, más que ofrecer felicidad, exigen autenticidad.

Tal vez la postura que acabo de señalar no pueda constituir una moralidad completa, ya que no tiene nada que decir, o nada suficiente que decir sobre la sociedad, y por tanto nada que decir precisamente sobre la vida de un hombre en su conjunto. Tal vez incluso, en la medida en que ello es así, descansa en una ilusión. Pero el hecho mismo de que exista y de que atraiga, exige una respuesta por parte de quien crea evidente que la felicidad general tiene que ser el foco de la moralidad; como la exige la moralidad religiosa (en el ejemplo extremo que hemos puesto del Protestantismo), en la medida en que su postura difiere radicalmente de una postura puramente secular. Pues aun suponiendo que su pretensión de trascendentalidad sea falsa, son seres humanos los que la han soñado, y necesitamos comprender por qué fue ese el contenido de su sueño. (El humanismo —en su sentido contemporáneo de movimiento secularista y antirreligioso— parece que rara vez se ha enfrentado plenamente con esta consecuencia inmediata de su propio punto de vista: que esta terrible realidad, la religión,

es una creación *humana.)* Pues los hombres, como cuestión de hecho, encuentran valores en cosas tales como la sumisión, la entrega, la incertidumbre, el riesgo, incluso en la desesperación y en el sufrimiento, y estos valores difícilmente pueden ponerse todos en relación con el ideal central de la *felicidad.* Y si encontramos algunas explicaciones, psicoanalíticas quizás, o incluso zoológicas, de esas actitudes, y llegamos a considerarlas como aberraciones que hay que tratar de evitar, entonces ciertamente estamos cambiando el mundo desde el punto de vista de una cierta moralidad, y no meramente haciendo el mundo más sensible a lo que la moral incuestionablemente es.

«Bienestar» era nuestro punto de partida; después nos hemos ocupado de su comprensión como «felicidad». Tal vez cabría decir que pese a que algunas clases de concepciones morales rechacen la felicidad como noción central, existe todavía una noción de bienestar, mucho más amplia y no por ello sin contenido, en relación con la cual no se produciría ese rechazo. Es una cuestión que tiene sentido, pero yo no sé cómo responderla. Por un lado, los casos más extremos parecen dejarnos con una noción de bienestar que realmente no se encuentra a mucha distancia de «ser como los hombres deben ser», y ahí no queda contenido alguno. Por otro lado, al caracterizar esas posiciones, uno habla de lo que los hombres, de hecho, encuentran valioso, o necesitan, o desean; y si alguien dijera —bastante oscuramente, por cierto— que los hombres necesitan un mundo en el que haya riesgo, incertidumbre y la posibilidad de desesperarse, entonces, de una moralidad que subrayara esto en contraste con las moralidades que quieren dejar todo lo más sosegado posible, habría que decir todavía que se refiere al bienestar de los hombres. Sin embargo, algo queda todavía excluido por el uso de este término: sistemas de valores o de preceptos que no prestan ninguna atención en absoluto a lo que podemos entender que los hombres necesitan o desean.

Utilitarismo

Al discutir la cuestión de si todas las concepciones morales se refieren en última instancia a la felicidad humana, no he supuesto que la cuestión sea la misma que la de si todas las concepciones morales tienen que ser en una u otra medida una versión del *utilitarismo*. Obviamente no se trata de la misma cuestión si tomamos el utilitarismo en el sentido estricto de que sólo hay un principio moral, el de buscar la mayor felicidad para el mayor número; que «felicidad» significa aquí placer y ausencia de dolor; y que este único principio moral (pues se trata de un principio moral) hay que aplicarlo a cada situación individual («utilitarismo del acto»). Obviamente, existen muchas formas en que una moralidad puede estar en última instancia referida a la felicidad humana sin por ello tener que identificarse con *eso*. Pero creo también que existen multitud de formas en que la moralidad puede en última instancia referirse a la felicidad humana sin identificarse con el utilitarismo, ni incluso tomado en un sentido más lato.

Una dificultad que surge al discutir este asunto es la falta de acuerdo en la cuestión de con qué extensión puede usarse el término «utilitarismo» sin que deje de ser adecuado o significativo. El término ha sido utilizado a veces para referirse a concepciones morales que nada en absoluto tienen que ver especialmente con la felicidad o con el placer; en este sentido se usa para referirse a cualquier concepción que sostiene que la rectitud o no rectitud de una acción depende siempre de las con-

secuencias de la acción, de su tendencia a conducir a estados de cosas intrínsecamente buenos o malos. Este sentido tan amplio —que probablemente quedaría reflejado mejor por el término «consecuencialismo» que por el término «utilitarismo»— no es el que vamos a discutir aquí; aquí nos interesan solamente las concepciones de esta clase que consideran expresamente la *felicidad* como la única cosa intrínsecamente buena, la que, se supone, tienen como objetivo las acciones y las instituciones sociales. Pero esa restricción todavía deja bastante espacio para diferentes clases de utilitarismo.

Discutir en el vacío qué es lo que podría o no podría contar como una forma reconocible de utilitarismo sería un ejercicio puramente verbal y sin interés. La cuestión solamente puede abordarse preguntando qué es lo *esencial* en la concepción utilitarista de la moralidad; y esto es algo que no puede descubrirse meramente, ni tampoco principalmente, consultando lo que tenían en la mente Bentham y J. S. Mill y otros exponentes clásicos del sistema, sino considerando cuáles son los atractivos de la concepción utilitarista para el pensamiento moral. A mi juicio son cuatro principalmente, lo que no significa negar que no se relacionen entre sí de una forma digna de examinarse. En primer lugar, es no-trascendental, y no hace apelación a nada fuera de la vida humana, y en particular no hace apelación a consideraciones religiosas. En particular, contribuye así a apoyar la demanda enteramente razonable de que la moralidad habría de quedar hoy, como cosa obvia, liberada del cristianismo. Puede incluso parecer que contribuye a apoyar —a causa de cierto conservadurismo que consideraré más tarde— una pretensión mucho menos razonable, y que Nietzsche se dio cuenta muy bien de que era una imbecilidad, que la moralidad así liberada del cristianismo habría de seguir siendo la misma que la antes vinculada al cristianismo. En manos más radicales, sin embargo, el utilitarismo promete cambios más radicales.

Segundo: su bien básico, la felicidad, parece mínimamente problemático: por muchas que sean las cosas en que las personas disienten, sin duda que todos por

lo menos desean ser felices, y tener como objetivo la mayor felicidad posible tiene sin duda que ser, cualesquiera sean las otras cosas que hayan de someterse a ésta, un fin razonable. Ahora bien, existe un notable problema en este punto sobre la transición del objetivo, supuestamente indiscutible, de buscar la propia felicidad, al objetivo, más discutible, de buscar la felicidad de otra persona, y el desafortunado Mill ha sido repetidas veces llamado al orden por los críticos por tratar (se dice) de llevar a cabo esa transición por medio de un argumento deductivo. Por mi parte dudo de que fuera esto lo que Mill trataba de hacer, pero en cualquiera de los casos el problema no tiene una fuerza especial contra el utilitarismo —no hay ninguna razón para suponer que el utilitarismo, más que ninguna otra teoría, tuviera que tener una fórmula mágica para sacar de su amoralismo al hombre amoral. Lo *esencial* es más bien que el utilitarismo, lo mismo en este que en otros aspectos, es una moralidad de *compromiso mínimo:* supuesto sólo el requisito mínimo para estar en el mundo moral, a saber, la voluntad de tomar en consideración los deseos de los otros tanto como los propios, el utilitarismo puede moverse ya en ese terreno. Una cuestión mucho más interesante es la de si el «indiscutible» objetivo de la felicidad puede de hecho ser puesto al servicio de los propósitos utilitaristas. En la sección anterior ya hemos visto alguna razón para dudar de que la felicidad tenga que ser el objetivo de la vida humana; pero incluso dejando de lado esas cuestiones, no está en absoluto claro que cualquier sentido en el que (más o menos) quepa admitir un fin tal como indiscutible, sea también un sentido en favor del cual pueda hacerse trabajar al utilitarismo. Esta es una cuestión central: estaremos en mejor situación para abordarla cuando hayamos considerado los atractivos tercero y cuarto del utilitarismo.

El tercer atractivo es que los asuntos morales pueden, en principio, decidirse por un cálculo empírico de las consecuencias. El pensamiento moral se hace empír· y, en cuestiones de gestión pública, se convierte en un asunto de ciencia social. Y esto siempre ha sido conside-

rado por muchos como uno de los rasgos más gratificantes del utilitarismo. No es que se piense que los cálculos son fáciles, o incluso prácticamente posibles en algunos casos; el encanto radica más bien en que la naturaleza de la dificultad pierde, por lo menos, todo su misterio.

En cuarto lugar, el utilitarismo proporciona una moneda corriente universal de pensamiento moral: los distintos intereses de las distintas partes, y las distintas clases de demandas que actúan sobre una de las partes, pueden todos (en principio) ser convertidos a la misma moneda, en términos de felicidad. Este supuesto, y esto es importante, tiene como consecuencia que una cierta clase de conflicto, bien conocido para algunas otras concepciones morales, es aquí imposible, a saber, el conflicto entre dos pretensiones que son válidas las dos, y sin embargo, irreconciliables. Bajo algún otro sistema, un hombre puede verse en una situación en la que (así le parece a él) cualquier cosa que haga implica que está haciendo algo malo. Para el utilitarismo esto es imposible. Las distintas exigencias a las que ese hombre puede sentirse sometido pueden ser reducidas todas a la medida común del principio de la mayor felicidad, y no puede haber otra idea coherente de la acción correcta o de la acción no correcta que la de *lo que es mejor hacer tomando todo en consideración:* y si dos formas de acción resultan iguales en este sentido, entonces realmente no importa por cuál de las dos decida uno. En contraposición con esto, muchas personas pueden pensar que una determinada forma de actuación era la mejor que podía adoptarse tomando todo en consideración en las circunstancias concretas del caso, pero que, no obstante, el actuar de esa forma implicaba hacer algo no correcto. Esta es una idea que, a mi juicio, en última instancia, al utilitarismo tiene que resultarle incoherente. Esta es una de las razones por las que puede decirse (y ciertamente con razón) que para el utilitarismo la tragedia es imposible; pero el utilitarismo tiene también consecuencias más amplias, si no más profundas, que esa.

El utilitarista podría ciertamente volverse sobre este tipo de ideas y, después de reflexionar un poco, invocar

cosas tales como las consecuencias socialmente deseables de que la gente sea un poco aprensiva en relación con ciertas acciones, aun cuando esas acciones, en las circunstancias del caso, sean las mejores que pueden emprenderse: más tarde volveremos sobre este tipo de argumentación. Pero lo que como utilitarista está obligado a hacer es considerar como objetivo indiscutible del pensamiento moral la reducción del conflicto, la eliminación, siempre que sea posible, de los conflictos de valores que no sean solubles sin secuelas. Aqui, como en cualquier otro sitio, su preocupación es la eficiencia: la generación del conflicto es un signo de ineficacia de un sistema de valores y el utilitarismo tiene un artificio general para eliminarlos o resolverlos. Pero alguien podría preguntarse si tal eficiencia es un objetivo indiscutible. Uno puede ciertamente reducir el conflicto, o hacer la vida más simple, reduciendo el rango de las demandas que está dispuesto a tomar en consideración; pero en ciertos casos esto podría ser considerado no como un triunfo de la racionalidad sino como una evasión cobarde, una negativa a ver lo que hay que ver (y aquí cabría preguntarse una vez más si el subjetivismo desactivado deja realmente todo donde estaba).

Por atrayente que pueda resultar en general el utilitarismo, este cuarto atractivo suyo puede ser puesto en tela de juicio por razones importantes. Otras dificultades se agolpan cuando uno considera lo que presupone. Pues sólo podremos ser capaces de utilizar el principio de la mayor felicidad como medida común de las pretensiones o demandas de todos y de cada uno, si la «felicidad» de que se habla es en algún sentido *comparable* y en algún sentido *aditiva*. Solamente si podemos comparar las felicidades aquí implicadas para las diferentes personas y sobre los diferentes resultados, y también juntarlos en alguna suerte de felicidad general, podremos hacer que la cosa funcione. En un plano técnico estos problemas han sido objeto de disciplinas tales como la economía del bienestar y la teoria de la preferencia, que han trabajado en el marco de supuestos muy artificiales y que sólo han tenido un moderado éxito a la hora de ela-

borar esos problemas para la teoría económica. Aquí nos ocupamos de dificultades más generales. Si la «felicidad» implicada ha de ser tal como para permitir al utilitarismo cumplir sus promesas tercera y cuarta, ¿puede ser también el indiscutible objetivo que se prometía en la segunda?

La respuesta a esto parece que es que no. Bentham ofreció una explicación de la felicidad en términos de placer y ausencia de dolor, de la que muy claramente se suponía que cumplía todas las promesas a la vez; pero aun cuando hubiera satisfecho (cosa que no hacía) las condiciones de ser calculable, comparable y aditiva, no cumplía la condición de ser un objetivo indiscutible: cuanto más parecía la clase de placer que concebiblemente podía ser tratado en esos términos cuasi matemáticos, tanto más perdía el aspecto de algo a lo que fuera evidente que cualquier hombre racional tenía que tender, cosa de la que, aunque a disgusto, se dio cuenta Mill. Si, por un lado, la concepción de la felicidad se hace lo suficientemente generosa como para incluir todo aquello a lo que pudiera razonablemente tenderse como a una vida satisfactoria o a un ingrediente de esa vida, entonces progresivamente va perdiendo el aspecto de algo que pudiera ajustarse a las condiciones tercera y cuarta. Aparte de otras cosas, existe la dificultad de que mucho de lo que la gente incluye de hecho en el contenido de una vida feliz son cosas que esencialmente implican otros valores, como son, por ejemplo, la integridad, o la espontaneidad, o la libertad, o el amor, o la autoexpresión artística: valores que no solamente no pueden ser tratados de la forma que las condiciones tercera y cuarta parecen imponer a la «felicidad» del utilitarismo, sino que además parece, por lo menos en el caso de algunos de ellos, que se da de hecho una contradicción en la tentativa de pensarlos como algo que podría ser tratado de esa manera.

Esta es, pues, la primera dificultad general con que se enfrenta el utilitarismo. Su «felicidad» tiene que satisfacer ciertas condiciones si es que se quiere mantener lo esencial del utilitarismo; y la condición de que indis-

cutiblemente habría de ser el objetivo de la aspiración humana entra en conflicto con las otras codiciones que tiene que satisfacer para poder ser tratada como el utilitarismo exige que se la trate. Enfrentado con esta dificultad general, una de las formas en que el utilitarismo tiende a reaccionar es la de poner en tela de juicio los valores implicados en las concepciones más intratables de la felicidad, como irracionales, tal vez, o como reminiscencias de una época pasada. Tales argumentos pueden ofrecer sin duda ciertos puntos interesantes, pero su estrategia global es vergonzosamente circular: la racionalidad utilitarista queda convertida en test de lo que ha de contar como felicidad, para quitar del paso toda otra clase de felicidad que puede constituir una objeción contra el utilitarismo. Todo lo que se necesita para enfrentarse a esto en el plano teórico es una firme negativa a sentirse intimidado.

El problema, sin embargo, no se reduce al nivel teórico: se presenta también, y drásticamente, en la sociedad, y la negativa a sentirse intimidado puede ser aquí inadecuada o difícil de hacerse valer. En los casos de planificación, conservación, bienestar, y decisiones sociales de todas clases, un conjunto de valores que, por lo menos en teoría, son cuantificados en términos de recursos, se enfrentan con valores que no son cuantificables en términos de recursos: tales como el valor de conservar una parte antigua de la ciudad, o de atender tanto a la dignidad como al confort de los pacientes en una unidad geriátrica. Una y otra vez, los defensores de tales valores se enfrentan con el dilema, o bien de negarse a cuantificar el valor en cuestión, en cuyo caso desaparece también por completo de la suma, o bien de tratar de asignarle alguna cantidad, en cuyo caso distorsionan el asunto de que se trata y por lo general pierden también la discusión, ya que el valor cuantificado no es suficiente para inclinar la balanza. En tales asuntos, no es que los utilitaristas estén comprometidos a pensar que esos otros valores no importan; o que estén confinados a considerar valiosas las cosas que de hecho son susceptibles de ser tratadas por medio del análisis coste-beneficio. Ni

tan siquiera, quizás, están obligados a pensar que todo valor social deba, en un caso dado, poder ser tratado en términos parecidos a los del análisis coste-beneficio: podrían decir que no tienen por qué suscribir la idea de que la medida común de la felicidad sea el dinero. Pero están obligados a sostener algo que en la práctica tiene esas implicaciones: que en última instancia no existen valores inconmensurables. Pues no es precisamente un rasgo accidental de la concepción utilitarista el que la presunción sea siempre en favor de lo cuantificable, y que los otros valores se vean por fuerza en el dilema apologético al que acabamos de referirnos. No es un accidente, porque (por una parte) el utilitarismo es, cosa nada sorprendente, el sistema de valores de una sociedad en la que los valores económicos son supremos; y también, en el plano teórico, porque la cuantificación en dinero es la forma obvia de aquello sobre lo que el utilitarismo insiste, la conmensurabilidad de los valores.

Existe una gran presión para la investigación de técnicas que hagan conmensurables esferas cada vez más amplias de valores sociales. Parte del esfuerzo debería dedicarse más bien a aprender —o tal vez, a reaprender— a pensar inteligentemente sobre conflictos de valores que son inconmensurables.

Me he referido a dificultades relativas al cumplimiento de las condiciones que el utilitarismo pone a la «felicidad». El hecho de que esas dificultades existan no significa en modo alguno que en un caso dado o en una determinada clase de casos no podamos tener una idea de cuál sería la solución utilitarista o de qué clases de cosas tendría en cuenta el utilitarista. Algo iría mal si esa conclusión fuera correcta, pues está claro que a veces podemos hacer esas cosas: al discutir el utilitarismo estamos discutiendo *algo*, y algo que con mucha frecuencia es perfectamente reconocible. Así pues, admitamos que, por lo menos en algunos casos, sabemos qué es lo que quiere decirse con analizar qué actuación conduciría a la mayor felicidad. Admitido esto, nos encontramos entonces con dos nuevas dificultades. Una es que el proceso de analizar tales consecuencias es, él

mismo, una actividad, una actividad que en las diversas circunstancias posee grados diversos de utilidad; y esto tiene que incluirse en la suma. La otra es que la respuesta a que se llega por el cálculo utilitarista del caso particular parece en ciertas ocasiones que es una respuesta moralmente incorrecta. Existe una clase de utilitarismo, denominado el utilitarismo de reglas *(rule-utilitarianism)*, cuyo objetivo es precisamente resolver estas dos dificultades al mismo tiempo, haciendo uso para ello de un mismo artificio.

El primer problema es que cualquier cálculo utilitarista efectivo habrá de tener lugar en condiciones de considerable incertidumbre y de información muy parcial, y así es probable que sus resultados no sean fiables. Además, el proceso de cálculo mismo lleva su tiempo; y el estar dispuesto a hacer cálculos en cada caso tiene rasgos psicológicos que, como cuestión de hecho, pueden ser un obstáculo para cosas que pueden ser deseables desde un punto de vista utilitarista, como puede ser una actuación rápida y decidida. Siendo así las cosas, cabe pensar que en la práctica podrían seguirse mejores consecuencias si los agentes, más que hacer cálculos para cada caso particular, suscribieran ciertas reglas que, sin cálculo alguno, se aplicaran usualmente a los casos particulares; sería la adopción de esas reglas y no la elección de acciones particulares lo que se examinaría recurriendo al principio de la mayor felicidad.

Esta misma idea se invoca para explicar el otro hecho, que, de otro modo, constituiría una dificultad, a saber: que fácilmente podemos imaginar casos —por ejemplo, el caso en el que la condena de un inocente fuera necesaria y suficiente para evitar grandes males— en los que el resultado utilitarista parece que entra en conflicto con lo que muchos considerarían que es la respuesta moralmente correcta; y al igual que la justicia en este caso, también el cumplimiento de las promesas y el decir la verdad presentan dificultades cuando se interpretan desde el punto de vista del utilitarismo del acto. Se espera que el utilitarismo de la regla pueda disolver estas dificultades pretendiendo que todo lo que tiene que demos-

trarse es que estas reglas o prácticas de justicia, de cumplimiento de las promesas o de veracidad poseen una utilidad positiva sobre sus alternativas.

Esto es sólo un bosquejo. Muchas cosas que difieren entre sí en aspectos importantes pueden caer bajo la denominación de «utilitarismo de reglas», y también son diferentes las cosas que hay que decir sobre ellas[5]. Todo lo que voy a tratar de hacer aquí va a ser sugerir un par de puntos en relación con la cuestión de hasta qué punto el utilitarista puede ir consistentemente en la dirección de las reglas; y sostener que, o bien no puede ir lo suficientemente lejos como para resolver la segunda dificultad, o si no, que tiene que ir tan lejos que (tanto el utilitarista como cualquier otro) dejará de ser utilitarista.

Es ciertamente posible para un utilitarista, sin incurrir por ello en incoherencia, adoptar una práctica general para tratar una determinada clase de casos, aun cuando algunas aplicaciones particulares de la práctica conduzcan a un resultado diferente del que se habría obtenido por medio de un cálculo individual de esos ejemplos. El paradigma de esto es el sistema de contabilidad que se aplica para muchos servicios públicos, sistema por el que, ocasionalmente, puede enviarse una factura de un dólar aun cuando procesar cada factura tiene un coste de dos dólares: la razón es que resulta más barato enviar todas las facturas en el momento debido, por pequeño que sea su importe, que interrumpir el procesamiento de todas para extraer unas cuantas. Vamos a llamar a esto el «modelo del recibo del gas».

Ahora bien, el modelo del recibo del gas se refiere a consecuencias de hecho: las consecuencias que tiene de hecho el limitarse a aplicar una regla, por un lado, y las consecuencias que de hecho tiene una interferencia particular, por otro. Esta clase de modelo, por lo menos, no puede hacer más digerible para un utilitarista coherente una forma de argumentación que no invoca ni

---

[5] Para una detallada y sutil discusión de estos asuntos, véase el importante libro de David Lyon, *The Forms and Limits of Utilitarianism*, Oxford, Clarendon Press, 1969.

las consecuencias que de hecho tiene una elección particular ni las consecuencias que de hecho tiene el que todos se atengan a una regla, sino las consecuencias *hipotéticas* del ajuste *imaginario* a una regla. Así, el modelo familiar de argumento moral consistente en la pregunta «¿Qué ocurriría si todos hicieran eso?» no puede tener ningún efecto sobre un utilitarista consistente, a menos que su acción hubiera de tener el efecto de que todo el mundo hiciera eso, cosa que por lo general es bastante implausible. Una consecuencia puramente imaginaria no podría figurar más en un cálculo utilitarista de lo que puede hacerlo la felicidad o infelicidad de personas puramente imaginarias. Así pues, el modelo del recibo del gas, por lo menos, no puede ponernos en el camino de esa clase de argumentación por generalización.

Si el utilitarista desea justificar el uso de una argumentación por generalización que se refiera a consecuencias imaginarias, tendrá que alejarse un paso más de las consecuencias que de hecho tienen las decisiones particulares, y tratar el asunto en términos de las consecuencias que *de hecho* tiene que la *gente piense en términos de consecuencias imaginadas*. Pero entonces parece que se aparta cada vez más de las ventajas iniciales del utilitarismo. Pues, primero, el supuesto cálculo de la utilidad de que la gente piense en términos de consecuencias imaginadas —más bien que examinando los casos particulares, o que examinando las consecuencias de reglas más específicas, o que considerando la moralidad local como un asunto de su incumbencia, o que muchas otras posibilidades—, ese cálculo empieza a tener cada vez más aspecto de ser engañoso en sus pretensiones. ¿Cómo podría saber el utilitarista cuáles podrían ser las consecuencias de esas diversas prácticas? En segundo lugar, sí que sabe una cosa por lo menos: que cuanto más general sea la práctica a la que se aplica el cálculo utilitarista, tantos más serán los casos en los que el cálculo particular para ese caso habría conducido a un resultado distinto, y con ello, tanto mayor también la inutilidad táctica que está permitiendo en su persecución de la uti-

lidad estratégica. Y en vistas de lo problemática que resulta esa utilidad estratégica, no tiene más remedio que preocuparle lo siguiente: una de las motivaciones del utilitarismo era, después de todo, un obstinado imperativo de pensar en términos de consecuencias calculables y no atenerse a la tradición, a las prácticas recibidas, etc.

Cuanto más considera uno el utilitarismo de reglas, tanto más urgentes resultan las cuestiones de esta clase. Volviendo una vez más al modelo del recibo del gas, recordemos que lo que principalmente hacía sensata esa práctica uniforme eran los costos que suponía el interferirla. Lo análogo de esto en la deliberación moral ordinaria es la inutilidad de calcular las consecuencias particulares. Pero el efecto del argumento queda cancelado si consideramos un caso en el que el cálculo particular *ha sido hecho ya;* y esto es así en los casos moralmente problemáticos que representaban la segunda clase de dificultades que se suponían quedaban solventadas por el utilitarismo de reglas. Si el cálculo se ha hecho ya, y ha resultado que las consecuencias de infringir la regla son mejores que las de respetarla, entonces las consideraciones sobre la inutilidad del cálculo no pueden cambiar este resultado. Y, desde luego, es difícil ver que, para un utilitarista coherente, pudiera haber algo, sea lo que ello fuere, capaz de cambiar este resultado. Cualquiera que sea la utilidad general de contar con una cierta regla, si uno ha alcanzado de hecho un punto desde el que ve que la utilidad de infringir la regla en una determinada ocasión es mayor que la de respetarla, entonces no cabe duda de que sería totalmente irracional no infringirla.

Ha habido, desde luego, utilitaristas duros que han sacado esa conclusión, como es el caso de J. J. C. Smart. Si al utilitarismo se le puede hacer funcionar en absoluto, entonces para mí está fuera de duda de que la actitud correcta ha de ser la de estos autores: el utilitarismo es una doctrina especial, que no tiene necesariamente que coincidir con las ideas morales contemporáneas de Occidente en todos los aspectos, y, siendo así, uno ha de esperar que tenga lo que muy bien podrían parecer con-

clusiones poco gratas. Contrariamente a esto, un rasgo de la moderna teoría utilitarista es que es una teoría sorprendentemente conformista. Bentham y Mill consideraron el principio de la mayor felicidad como un instrumento de crítica, y pensaban que, valiéndose de él, podrían mostrar que buena parte de las creencias morales de la época victoriana eran erróneas e irracionales, como por supuesto lo eran. Pero, con excepción de las bien delimitadas áreas de la reforma sexual y de la reforma penal, cuyo tratamiento es el heredado de Bentham y Mill, la teoría utilitarista moderna tiende a gastar más esfuerzo en reconciliar el utilitarismo con las creencias morales existentes que en rechazar esas creencias apoyándose en el utilitarismo. Un autor reciente, por ejemplo, se ha tomado un honesto y gran esfuerzo en mostrar que las ejecuciones públicas no podrían justificarse, como podría parecer, con argumentos utilitaristas. Nuestro autor se queda con algunas francas dudas; pero éstas son dudas sobre la aplicación y la formulación del utilitarismo, y no, como deberían ser, dudas sobre si no podrian ser reintroducidas las ejecuciones públicas. Este es un caso absurdo. Pero en términos más generales, todas las cualidades humanas que son consideradas valiosas y que, sin embargo, se resisten a un tratamiento utilitarista, como son la intransigente pasión por la justicia, ciertas clases de coraje, la espontaneidad, la disposición a oponerse a cosas tales como los experimentos útiles sobre pacientes ancianos o las bombas de napalm sobre ciertas gentes para asegurar (se dice) la felicidad de más gente, generan con frecuencia en los teóricos del utilitarismo el deseo de acomodar el utilitarismo a esos valores, y no la condena de tales valores como un legado irracional de la era preutilitarista. Esto es sin duda un tributo a la honestidad e imaginación de esos utilitaristas, pero no a su consistencia o a su utilitarismo.

El utilitarismo de reglas, como la empresa que es de tratar de atenerse a algo distintivamente utilitarista, procurando al mismo tiempo quitarle sus puntas más hirientes, me parece un fracaso. Ese terreno intermedio no es habitable lógicamente. Frente a él, uno podría,

por un lado, adoptar la línea a la que ya me he referido, la de Smart y otros, y seguir con el utilitarismo de actos modificado solamente dentro de los límites reconocibles del modelo del recibo del gas. Esto es, por lo menos, coherente con los equivocados fines del utilitarismo, y el hecho de que arroje algunos resultados morales particulares distintivos y (para muchos) indigeribles no debería ser motivo de sorpresa. Si, por otro lado, se abandona este territorio, y se empieza a aplicar el principio utilitarista a prácticas y a hábitos de pensamiento cada vez más generales, es muy improbable que lo que que quede tenga algún contenido que pueda considederarse distintivamente utilitarista. Esta capacidad del utilitarismo para aniquilarse a sí mismo, una vez abandonado el nivel básico, puede quedar ilustrada con un breve argumento, con el que voy a terminar. Sus premisas empíricas quizás no estén fuera de toda duda. Pero sí que son por lo menos tan plausibles como las que los utilitaristas suelen utilizar en estas materias.

Uno de los efectos perturbadores cuando la gente se convierte en utilitaristas activos y conscientes es que tienden a degradar el patrón de medida moral: comienza a operar una ley de Gresham por la que las malas acciones de los hombres malos obliga a hombres mejores a hacer cosas que, en mejores circunstancias, serían también malas. La razón de esto es simple: un utilitarista está siempre justificado a hacer la cosa menos mala que sea necesaria para evitar la cosa peor, cosa peor que, si no lo impidiera la menos mala, ocurriría en las circunstancias dadas (incluyendo, por supuesto, la cosa peor que cualquier otro podría hacer) y lo que de esta forma está justificado a hacer puede ser algo que, considerado en sí mismo, es bastante sucio. El acto de evitación forma parte de las concepciones utilitaristas, y ciertas nociones de responsabilidad negativa (el que uno sea responsable de lo que deja de evitar tanto como de lo que hace) son, por lo mismo, también características del utilitarismo. Si esto es así, es empíricamente probable que pueda esperarse una escalada de la actividad de evitación; y las consecuencias totales de esto, *medidas en los propios*

estándares utilitaristas, serán peores que si nunca se hubiera iniciado esa escalada.

Sin embargo, el utilitarista que está inmerso en el sistema no puede hacer nada para evitar esto; tiene que pensar en términos de las consecuencias fácticas en curso, y ningún amago de principio podrá hacer ahora nada (o por lo menos, nada útil) en el terreno de esas consecuencias fácticas en curso —desde donde está no hay forma de dirigir una incursión a un terreno moralmente más alto. Volviendo, sin embargo, sobre sí mismo, puede considerar cómo los objetivos utilitaristas podrían haber sido realizados mejor de lo que lo han sido en un mundo de utilitaristas mezclados con villanos. Y sin duda que hubieran sido mejor realizados si no hubiera habido villanos: pero esto es ciertamente utópico. Lo más prometedor parece que es un estado de cosas en que un número suficiente de personas se resista a proseguir esta carrera de envilecimiento: que se resistan, por ejemplo, fijando una serie de cosas que no puede entrar en consideración hacer, o que no están dispuestos ellos a hacer o que no pueden soportar que se hagan, cualesquiera sean las cosas que los otros hacen o pueden hacer. Existe un límite a sus actividades de evitación. Y para ello parece que suficiente gente, durante un tiempo suficiente, tiene que estar dispuesta a obstinarse en hacer varias cosas, cualesquiera que sean las consecuencias. Esto significa que suficiente gente, durante un tiempo suficiente, no han de pensar como utilitaristas; sin duda alguna, han de pensar como no utilitaristas. Y no podrán preservar en la trastienda de sus mentes la actitud utilitarista en coexistencia con la requerida obstinación moral. Pues tienen que ser capaces de resistir la tentación utilitaria en las circunstancias más difíciles, es decir, cuando es obvio que de resistirla se van a seguir muchos males, y para esto su no-utilitarismo tiene que estar profundamente enraizado.

Algunos utilitaristas han llegado, aunque no exactamente por las mismas razones, a algo parecido a esta conclusión y han pensado que lo que con ello se demostraba es que de la verdad del utilitarismo podría estar

al tanto una élite responsable, pero que no debería estar excesivamente divulgada entre las masas. Esta propuesta, tanto desde un punto de vista personal como desde un punto de vista social, es una propuesta sin perspectivas. Desde un punto de vista personal, porque el estado mental que se atribuye al utilitarista reflexivo, y la actitud hacia otro que ese estado implica, sólo podrían ser honestamente mantenidas, si es que pueden serlo, por un hombre muy inocente (tal vez Sidgwick lo era), y ningún hombre reflexivo de nuestra época puede ser ese inocente. Desde un punto de vista social, porque las instituciones educativas y otras instituciones que se requerirían para encarnar esa idea tendrían que ser muy diferentes de todo lo que ahora podemos esperar o tolerar, o de todo lo que el propio utilitarismo podría desear.

Si todo esto es verdad, entonces el mundo que el utilitarista reflexivo tiene finalmente que escoger, como el mundo que más probablemente puede arrojar los resultados que el utilitarista desea, es un mundo en el que la ley de Gresham no rige porque un número suficiente de personas durante un tiempo suficiente se muestran profundamente recalcitrantes a pensar en términos utilitaristas. No es posible que esta disposición pudiera coexistir con la creencia en el utilitarismo; ni tampoco parece aceptable o socialmente posible que la mayoría tuviera que tener esta disposición, mientras que otros, la élite utilitarista, creyeran en el utilitarismo. Todo lo que queda es que el mundo que satisfaría las aspiraciones del utilitarista sería un mundo del que estuviera totalmente ausente la creencia en el utilitarismo como una doctrina moral omnicomprensiva, excepto tal vez como una excentricidad menor y sin trascendencia.

De este modo, si el utilitarismo es verdadero, y si también son verdaderas algunas premisas empíricas bastante plausibles, entonces es mejor que la gente no crea en el utilitarismo. Si por otro lado, el utilitarismo es falso, entonces por supuesto que es mejor que la gente no crea en él. Así pues, de cualquier forma es mejor que la gente no crea en él.